KATHOLIEKE UNIVERSITEIT LEUVEN

Faculty of Economics and Applied Economics

Department of Economics

The revealed preference approach to collective consumption behavior: testing, recovery and welfare analysis

by

Laurens CHERCHYE
Bram DE ROCK
Frederic VERMEULEN

Public Economics

Center for Economic Studies
Discussions Paper Series (DPS) 07.24
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

September 2007

DISCUSSION PAPER

The revealed preference approach to collective consumption behavior: testing, recovery and welfare analysis

Laurens Cherchye*, Bram De Rock† and Frederic Vermeulen‡

September, 2007

Abstract

We extend the nonparametric 'revealed preference' methodology for analyzing collective consumption behavior (with consumption externalities and public consumption), to render it useful for empirical applications that deal with welfare-related questions. First, we provide a nonparametric necessary and sufficient condition for collectively rational group behavior that incorporates the possibility of assignable quantity information. This characterizes collective rationality in terms of feasible personalized prices, personalized quantities and income shares (representing the underlying sharing rule). Subsequently, we present nonparametric testing tools for data consistency with special cases of the collective model, which impose specific structure on the preferences of the group members (in terms of consumption externalities and public consumption); and we show that these testing tools in turn allow for nonparametrically recovering (bounds on) feasible personalized prices, personalized quantities and income shares that underlie observed (collectively rational) group behavior. In addition, we present formally similar testing and recovery tools for the general collective consumption model, which imposes minimal *a priori* structure. Interestingly, the proposed testing and recovery methodology can be implemented through integer programming (IP and MILP), which is attractive for practical applications. Finally, while we argue that assignable quantity information generally entails more powerful recovery results, we also demonstrate that precise nonparametric recovery (i.e. tight bounds) can be obtained even if no assignable quantity information is available.

Keywords: collective model, revealed preferences, consumption, testing, recovery, welfare analysis, integer programming.

*University of Leuven, Campus Kortrijk and Center for Economic Studies; Fund for Scientific Research - Flanders (FWO-Vlaanderen). E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven-kortrijk.be.

†University of Leuven, Campus Kortrijk and Center for Economic Studies. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: bram.derock@kuleuven-kortrijk.be.

‡CentER, Department of Econometrics and OR, Tilburg University and Netspar. P.O. Box 90153, NL-5000 LE Tilburg, The Netherlands. E-mail: frederic.vermeulen@uvt.nl.

1. Introduction

The 'collective' consumption model explicitly recognizes that group (e.g. household) consumption is the outcome of multi-person decision making, with each individual decision maker (e.g. household member) characterized by her or his own rational preferences. Following Chiappori (1988, 1992), it regards 'rational' group consumption as the Pareto efficient outcome of a within-group bargaining process. This collective approach contrasts with the conventional 'unitary' approach, which models groups as if they were single decision makers. The fact that the collective approach starts from individual preferences (and not 'group preferences') makes it particularly useful for addressing welfare-related questions that specifically focus on the within-group distribution of the group income.

For example, the 'targeting view' of Blundell, Chiappori and Meghir (2005) takes as a starting point that the effectiveness of a specific benefit or tax also depends on the particular group (e.g. household) member to whom it has been targeted; and these authors argue that a unitary set-up, which implicitly assumes income pooling at the aggregate group level, fails to adequately deal with such targeting considerations. In addition, the collective model allows for analyzing welfare at the individual group member level rather than at the aggregate group level; for example, Browning, Chiappori and Lewbel (2006) suggest a collective approach for comparing the cost-of-living of individuals living alone with that of the same individuals living in a multi-member household. Finally, a concept that is intrinsically related to the collective approach is the so-called 'sharing rule', which divides the aggregate group means over the individual group members. Recovering this sharing rule, and subsequently explaining its variation in terms of group (member) characteristics, can yield useful insights into the distribution of the within-group bargaining power across the individual group members; see, for example, Browning, Bourguignon, Chiappori and Lechene (1994), Browning and Chiappori (1998) and Chiappori and Ekeland (2006).

Cherchye, De Rock and Vermeulen (2007) recently established a nonparametric 'revealed preference' characterization of a collective consumption model that considers general preferences of the individual group members, which allow for consumption externalities and public consumption within the group.[1] They introduced a testable necessary condition and a testable sufficient condition for data consistency with the collective consumption model that only require price and quantity data pertaining to the aggregate group level; these conditions have a similar formal structure as the *generalized axiom of revealed preference* (*GARP*) condition for the unitary model (Varian, 1982, building on Afriat, 1967). These nonparametric conditions allow for *testing* consistency of observed group behavior with collective rationality. By contrast, the empirical analysis of the welfare-related questions listed above requires *recovery* of the decision structure underlying the observed (aggregate) group behavior.

This paper focuses on the nonparametric analysis of recovery questions that are relevant for the collective model. More specifically, we explore whether the structural collective consump-

[1] Browning and Chiappori (1998) originally suggested this collective consumption model, and established its parametric characterization; see also Chiappori and Ekeland (2006) for additional discussion. Browning, Chiappori and Lewbel (2006) recently proposed a collective consumption model that explicitly accounts for economies of scale within the process of household consumption. Our model (implicitly) includes such economies of scale that follow from public consumption and consumption externalities.

tion model (i.e. individual preferences, individual consumption and the sharing rule) can be recovered on the basis of observed group behavior alone (i.e. aggregate quantities and prices). The recovery methodology that we present consequently enables the nonparametric analysis of welfare-analytical questions that are specific to the collective consumption model. Nonparametric recovery typically aims at identifying the *set* of structural models that are consistent with a given set of observations. The corresponding recovery questions are essentially the nonparametric counterparts of the so-called 'identification' questions in the parametric literature; see Chiappori and Ekeland (2005) for a general discussion on parametric identification for the collective model. To illustrate the difference between parametric and nonparametric recovery/identification, let us consider the unitary model. For that model, parametric identification aims at recovering the (structural model) parameters of a *pre-specified* utility function representing *unique* preferences from a set of demand (reduced form) parameters that are estimated. By contrast, from a nonparametric perspective, there usually are *many* preferences that are consistent with the same set of data satisfying the unitary *GARP* condition. Therefore, nonparametric recovery of the unitary model focuses on identifying the *set* of preferences that are consistent with a given data set; see, for example, Afriat (1967) and Varian (1982 and 2006). The main purpose of the current paper is to develop similar 'set identification' results for the collective model. In fact, given that this collective model includes the unitary model as a special case (i.e. when there is a single group member/decision maker), we also complement the existing literature on nonparametric recovery within the context of the unitary model.

In what follows, we will make the distinction between 'special cases' and 'the general case' of the collective model. The 'special cases' impose specific *a priori* structure on the group behavior: (1) the case in which all goods are publicly consumed, and (2) the case in which all goods are privately consumed and there are no consumption externalities. For these cases, we establish conditions that can be tested on the basis of the available (aggregate) price and quantity data and that are simultaneously necessary and sufficient for data consistency with collective rationality. In turn, these necessary and sufficient conditions allow for 'full' nonparametric recovery of the collective model. Interestingly, these results comply with those of Chiappori and Ekeland (2005), who consider similar special cases to obtain identification (or more precisely 'identifiability') within a parametric context. As we will discuss, our treatment of these special cases allows for a number of useful extensions, such as: nonparametric testing and recovery for the 'hybrid' case in which some goods are publicly consumed while all other goods are privately consumed without externalities; forecasting collective consumption behavior in new situations; and testing specific hypotheses regarding the collective decision process. Next, 'the general case' does not impose *a priori* structure and thus allows for public consumption and externalities of any good. For this case, we develop a necessary condition for collective rationality that can be tested on the available data. And we subsequently argue that this necessary condition provides a useful basis for nonparametric recovery of the sharing rule underlying observed group behavior.

The rest of the paper unfolds as follows. Section 2 recaptures the nonparametric condition for collective rationality. We extend the discussion of Cherchye, De Rock and Vermeulen (2007) by including the possible use of 'assignable quantity' information (which is often partly, but not fully, available in budget surveys). This characterizes collectively rational group behavior in terms of 'feasible personalized prices, personalized quantities and income shares' (representing the underlying sharing rule). Section 3 considers the 'special cases' mentioned above. For these cases, we develop (necessary and sufficient) tests for data consistency with collectively rational behavior that merely involve mixed integer linear programming (MILP), which is attractive for practical applications. Subsequently, we address the recovery issue and discuss the possibility to identify (through MILP) the sets, and corresponding upper and lower bounds, of feasible personalized prices, personalized quantities and income shares that are consistent with collectively rational group behavior; these results parallel the results on nonparametric preference recovery for the unitary model. Additional assignable quantity information generally entails tighter bounds; but, as we will show, precise recovery (i.e. tight bounds) can be obtained even if no assignable quantity information is available. Section 4 addresses the testing issue for 'the general case'; we introduce a (necessary) test for data consistency with collectively rational consumption behavior which can be implemented through integer programming (IP). Section 5 considers the corresponding recovery issue, and demonstrates the possibility to identify (through MILP) bounds on the feasible income shares without imposing specific *a priori* structure on the collective model. Still, recovery of feasible personalized quantities and prices remains impossible for the general case, which again falls in line with the results of Chiappori and Ekeland (2005) on parametric identification. Section 6 summarizes and offers some concluding remarks regarding the practical application of the proposed methodology. The appendix contains the proofs of our main results.

2. Rational collective consumption behavior

This section introduces the nonparametric characterization of the collective model that considers general preferences of the group members, which impose minimal *a priori* structure on the consumption externalities and public consumption within the group. Starting from this nonparametric characterization, Section 3 addresses testing and recovery of special cases of this general model, which include additional structure on the nature of the members' preferences. Sections 4 and 5 subsequently return to (testing and recovery of) the general model that is presented here.

2.1. Individual preferences

We consider an M-member group. The group purchases the (non-zero) n-vector of quantities $\mathbf{q} \in \mathbb{R}^n_+$ with corresponding prices $\mathbf{p} \in \mathbb{R}^n_{++}$. All goods can be consumed privately, publicly or both. For example, car use may be partly public (e.g. car use for a family trip) and partly private (e.g. car use for work). In addition, as for the privately consumed quantities, we allow for externalities (which includes the possibility of 'altruism'). Summarizing, this obtains

$$\mathbf{q} = \sum_{m=1}^{M} \mathbf{q}^m + \left(\sum_{m=1}^{M} \mathbf{Q}^m + \mathbf{Q}^h \right)$$

with $\mathbf{q}^m \in \mathbb{R}^n_+$ the private consumption quantities of member m *without* externalities (i.e. that *do not* enter the utility function of at least one other member), $\mathbf{Q}^m \in \mathbb{R}^n_+$ the private consumption quantities of member m *with* externalities (i.e. that *do* enter other members' utility functions), and $\mathbf{Q}^h \in \mathbb{R}^n_+$ the publicly consumed quantities.

Note that not only the quantities $\mathbf{Q}^h$ but also the quantities $\mathbf{Q}^m$ may be interpreted as 'public consumption', given that they enter other members' utility functions. To simplify notation, we therefore use $\mathbf{Q} = \left(\mathbf{Q}^1, ..., \mathbf{Q}^M, \mathbf{Q}^h\right) \in \left(\mathbb{R}^n_+\right)^{M+1}$ in the following. No qualitative distinction can be made between the different components of $\mathbf{Q}$. Yet, there is a clear quantitative difference: group members may accord another marginal valuation to private consumption $\mathbf{Q}^m$ than to public consumption $\mathbf{Q}^h$. Our use of the simplified notation $\mathbf{Q}$ rather than $\left(\mathbf{Q}^1, ..., \mathbf{Q}^M, \mathbf{Q}^h\right)$ also falls in line with the argument of Chiappori and Ekeland (2006), who state that privateness (versus publicness) of consumption has no testable implication *per se* if no additional information (on 'assignable quantities'; see below) is used.

Formally, we assume that preferences of each member m can be represented by a non-satiated utility function $U^m(\mathbf{q}^m, \mathbf{Q})$ that is non-decreasing in its arguments. Given the construction of $\mathbf{Q}$, this effectively accounts for public consumption within the group and (positive) consumption externalities.

2.2. Assignable quantities

We start from T observations of group consumption quantities under different price regimes. For each observation t we use $\mathbf{p}_t$ and $\mathbf{q}_t$ to denote the observed prices and aggregate quantities. In general, for each $\mathbf{q}_t$ we do not observe its constituent components $\mathbf{q}_t^m$ and $\mathbf{Q}_t$. If we observe how much a group member consumes of a particular good, then we say this good is 'assignable'; see Bourguignon, Browning and Chiappori (2006). In this paper, we consider *assignable quantities* that specifically relate to private quantities without externalities ($\mathbf{q}_t^m$); these quantities can be assigned to a single member's utility function, which is no longer the case if externalities are possible. Of course, in practice such assignable quantities necessarily involve an assumption that preference externalities are effectively absent. Note further that it may well be that we have such assignable quantity information for only a subset of group members rather than for all members. Chiappori and Ekeland (2005), considering general collective consumption models, which also includes private consumption, argue that such assignable quantity information (on $\mathbf{q}_t^m$) is necessary for obtaining 'identifiability'. More specifically, they show that it is necessary for parametrically recovering the underlying structure of the consumption model (i.e. member preferences and decision process) from the group's aggregate consumption behavior alone. In the following, we will argue that the use of assignable quantity information can (often considerably) enhance the power of the nonparametric analysis. Still, we will also demonstrate that precise nonparametric recovery is sometimes possible even if no assignable quantity information is available.

For each observation t, we define the (observed) *assignable quantities* $\mathbf{q}_t^{Am} \in \mathbb{R}^n_+$ for member m as lower bounds for the (unobserved) quantities $\mathbf{q}_t^m$, i.e.

$$\mathbf{q}_t^m \geq \mathbf{q}_t^{Am}.$$

Our following discussion focuses on a set of observations $S^A = \{(\mathbf{p}_t,\ \mathbf{q}_t;\ \mathbf{q}_t^{A1}, ...,\ \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$. The superscript A in S^A refers to the fact that this set includes assignable quantities. We note that, in our general case, for some goods it may well be that only *parts* of the consumed quantities are assignable (e.g., car use for work can be assignable while car use for a family trip is clearly not).

Let us consider some specific examples. For simplicity, we focus on two-member households ($M = 2$) consisting of a wife (member 1) and a husband (member 2):

1. A first example implies that all goods are fully assignable, which means $\mathbf{q}_t = \mathbf{q}_t^{A1} + \mathbf{q}_t^{A2}$ and thus $\mathbf{q}_t^m = \mathbf{q}_t^{Am}$. For example, Bonke and Browning (2006) discuss a data set on household consumption that could be used in this case. Importantly, given our specific assumption of assignable quantities (which -to recall- pertains to private consumption without externalities), such a full assignability assumption excludes public consumption and consumption externalities (because $\mathbf{Q}_t = (\mathbf{0}, ..., \mathbf{0}, \mathbf{0})$); i.e. group members are of the so-called 'egoistic' type.

   A specific application of this example setting includes an observation t of member m's consumption behavior when deciding alone (rather than in group): e.g., for the wife as member 1, this corresponds to $\mathbf{q}_t^{A1} = \mathbf{q}_t$ if the full consumption quantity can be assigned to the wife in situation t. Note that such an application implies that we assume that individual egoistic preferences do not change when deciding in group (e.g. when living in a multi-member household) or when deciding alone (e.g. when living apart); for instance, one may assume constant preferences for the wife (husband) in a couple and the same wife (husband) as a widow(er) (compare with Michaud and Vermeulen, 2006). In fact, the testing tools presented below effectively allow for *testing* such a constant preference assumption.

2. Our general set-up also includes intermediate scenarios with $\mathbf{q}_t^{Am} \neq \mathbf{q}_t^m$ and $\mathbf{q}_t^{Am} \neq \mathbf{0}$. Generally, this intermediate case includes settings characterized by assignable goods as well as non-assignable goods, which can be characterized by externalities as well as public consumption. For instance, a model that is often considered in the literature excludes, like before, public consumption and consumption externalities ($\mathbf{Q}_t = (\mathbf{0}, ..., \mathbf{0}, \mathbf{0})$) while, different from before, it *only* uses quantity information on a so-called 'exclusive good' for each household member (i.e. an assignable good that is exclusively consumed by the wife or the husband; see Bourguignon, Browning and Chiappori, 2006). A specific application is Chiappori's (1988) labor supply model with egoistic household members; in that setting, each household member's leisure is the exclusive good while the other, Hicksian consumption good is non-assignable.

3. A final example implies no assignable quantity information, i.e. $\mathbf{q}_t^{Am} = \mathbf{0}$. In that case, there are no restrictions on $\mathbf{q}_t^m$ and $\mathbf{Q}_t$ apart from non-negativity and adding-up ($\mathbf{q}_t = \sum_{m=1}^M \mathbf{q}_t^m + \left(\sum_{m=1}^M \mathbf{Q}_t^m + \mathbf{Q}_t^h\right)$). This includes the setting in which all goods, even if assignable, can be characterized by externalities. Cherchye, De Rock and Vermeulen (2007) established nonparametric empirical restrictions for collectively rational group behavior in this scenario, which imposes minimal *a priori* restrictions.

2.3. Collective rationality

To define the collective rationality condition, we focus on feasible decompositions of the aggregate quantities $\mathbf{q}_t$ in terms of $\mathbf{q}_t^m$, the private quantities that only enter member m's utility function, and $\mathbf{Q}_t$, the private and public quantities that enter other members' utility functions. Specifically, we define *feasible personalized quantities* $\widehat{\mathbf{q}}_t$, which capture such feasible decompositions of $\mathbf{q}_t$.

Definition 1. *Let* $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ *be a set of observations. For each observation* t*,* feasible personalized quantities $\widehat{\mathbf{q}}_t = \left(\mathfrak{q}_t^1, ..., \mathfrak{q}_t^M, \mathfrak{Q}_t\right)$ *satisfy* $\mathfrak{q}_t^m \geq \mathbf{q}_t^{Am}$*,* $m = 1, ..., M$*, and* $\mathfrak{Q}_t = \left(\mathfrak{Q}_t^1, ..., \mathfrak{Q}_t^M, \mathfrak{Q}_t^h\right) \in \left(\mathbb{R}_+^n\right)^{M+1}$ *such that* $\mathbf{q}_t = \sum_{m=1}^M \mathfrak{q}_t^m + \left(\sum_{m=1}^M \mathfrak{Q}_t^m + \mathfrak{Q}_t^h\right)$.

Example 1 illustrates the concept.

Example 1. *Consider a two-member household* ($M = 2$) *with a wife (member* 1*) and a husband (member* 2*) that consumes three goods* ($n = 3$)*. Suppose two observations with aggregate quantities*

$$\mathbf{q}_1 = (3, 5, 4)' \text{ and } \mathbf{q}_2 = (4, 3, 5)',$$

and assignable quantities

$$\begin{aligned} \mathbf{q}_1^{A1} &= (0, 0, 1)' \text{ and } \mathbf{q}_1^{A2} = (0, 2, 0)'; \\ \mathbf{q}_2^{A1} &= (1, 0, 0)' \text{ and } \mathbf{q}_2^{A2} = (0, 0, 2)'. \end{aligned}$$

One possible specification of the feasible personalized quantities $\widehat{\mathbf{q}}_1$ *and* $\widehat{\mathbf{q}}_2$ *is then*

$$\begin{aligned} \mathfrak{q}_1^1 &= (2, 0, 1)', \mathfrak{q}_1^2 = (1, 2, 0)', \mathfrak{Q}_1^1 = (0, 0, 0)', \mathfrak{Q}_1^2 = (0, 0, 3)', \mathfrak{Q}_1^h = (0, 3, 0)'; \\ \mathfrak{q}_2^1 &= (1, 2, 0)', \mathfrak{q}_2^2 = (0, 1, 2)', \mathfrak{Q}_2^1 = (3, 0, 0)', \mathfrak{Q}_2^2 = (0, 0, 0)', \mathfrak{Q}_2^h = (0, 0, 3)'. \end{aligned}$$

Using the concept of feasible personalized quantities, we can define the condition for a collective rationalization of a set of observations S^A, which basically requires that the observed group consumption can be represented as a Pareto efficient outcome of some within-group bargaining process.

Definition 2. *Let* $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ *be a set of observations. A combination of* M *utility functions* $U^1, ..., U^M$ *provides a* collective rationalization *of* S^A *if for each observation* t *there exist feasible personalized quantities* $\widehat{\mathbf{q}}_t = \left(\mathfrak{q}_t^1, ..., \mathfrak{q}_t^M, \mathfrak{Q}_t\right)$ *and* $\mu_t^m \in \mathbb{R}_{++}$*,* $m = 1, ..., M$*, such that*

$$\sum_{m=1}^M \mu_t^m U^m\left(\mathfrak{q}_t^m, \mathfrak{Q}_t\right) \geq \sum_{m=1}^M \mu_t^m U^m\left(\mathfrak{z}^m, \mathfrak{Z}^1, ..., \mathfrak{Z}^M, \mathfrak{Z}^h\right)$$

for all $\mathfrak{z}^m, \mathfrak{Z}^1, ..., \mathfrak{Z}^M, \mathfrak{Z}^h \in \mathbb{R}_+^n$ *with* $\mathbf{p}_t'[\sum_{m=1}^M \mathfrak{z}^m + (\sum_{m=1}^M \mathfrak{Z}^m + \mathfrak{Z}^h)] \leq \mathbf{p}_t'\mathbf{q}_t$ *and* $\mathfrak{z}^m \geq \mathbf{q}_t^{Am}$.

Thus, a collective rationalization of S^A requires that there exists, for each observation t with assignable quantities $\mathbf{q}_t^{Am}$, feasible personalized quantities $\widehat{\mathbf{q}}_t$ that maximize a weighted sum of the group members' utilities U^m for the given group budget $\mathbf{p}_t'\mathbf{q}_t$. This optimality condition reflects the Pareto efficiency assumption regarding observed group consumption in the collective model. Each weight μ_t^m represents the 'bargaining power' of member m in observation t. See also Browning and Chiappori (1998) for a detailed discussion.

Clearly, assignable quantity information restricts the feasible set of utility functions U^m and bargaining weights μ_t^m in Definition 2. And thus, intuitively, additional assignable quantity information will yield more stringent nonparametric conditions for collective rationality. In turn, these stronger conditions will entail 'more powerful' nonparametric recovery results. We will repeatedly illustrate this in the sequel.

2.4. Nonparametric condition

We next establish a nonparametric condition for a collective rationalization of a set S^A. To do so, we first define *feasible personalized prices* $\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M\right)$ for observed aggregate prices $\mathbf{p}_t$, which complement the concept of feasible personalized quantities in Definition 1. We use $\widehat{\mathbf{p}}_t^m = \left(\mathfrak{p}_t^{m,1}, ..., \mathfrak{p}_t^{m,M}, \mathfrak{P}_t^m\right)$; and the interpretation of the different components is as follows. As for the first M components, the personalized prices equal the observed prices for member m's own private consumption quantities without externalities (i.e. $\mathfrak{p}_t^{m,m} = \mathbf{p}_t$ for the quantities $\mathfrak{q}_t^m$), while they equal zero for the other members' private consumption quantities without externalities (i.e. $\mathfrak{p}_t^{m,l} = \mathbf{0}$ for the quantities $\mathfrak{q}_t^l$, $l \neq m$). The remaining component $\mathfrak{P}_t^m = \left(\mathfrak{P}_t^{m,1}, ..., \mathfrak{P}_t^{m,M}, \mathfrak{P}_t^{m,h}\right)$ captures the fraction of the price for the quantities $\mathfrak{Q}_t$ that is borne by member m: for each separate component of $\mathfrak{Q}_t$ the corresponding personalized prices can be interpreted as Lindahl prices and must add up to the observed prices. More specifically, feasible personalized prices $\mathfrak{P}_t^{m,l}$, $l = 1, ..., M$, pertain to private quantities with externalities and feasible personalized prices $\mathfrak{P}_t^{m,h}$ to public quantities. Summarizing, we get the following formal definition.

Definition 3. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ be a set of observations. For each observation t,* feasible personalized prices $\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M\right)$ *with $\widehat{\mathbf{p}}_t^m = \left(\mathfrak{p}_t^{m,1}, ..., \mathfrak{p}_t^{m,M}, \mathfrak{P}_t^m\right)$, $m = 1, ..., M$, satisfy $\mathfrak{p}_t^{m,m} = \mathbf{p}_t$, $\mathfrak{p}_t^{m,l} = \mathbf{0}$ for $l \neq m$ and $\mathfrak{P}_t^m = \left(\mathfrak{P}_t^{m,1}, ..., \mathfrak{P}_t^{m,M}, \mathfrak{P}_t^{m,h}\right) \in \left(\mathbb{R}_+^n\right)^{M+1}$ such that $\mathbf{p}_t = \sum_{m=1}^M \mathfrak{P}_t^{m,c}$ for $c = 1, ..., M, h$.*

Example 2 illustrates the concept.

Example 2. *We recapture the situation of Example 1. Suppose the corresponding observed prices*

$$\mathbf{p}_1 = (1,3,2)' \text{ and } \mathbf{p}_2 = (2,1,3)'.$$

One possible specification of the feasible personalized prices $(\widehat{\mathbf{p}}_t^1, \widehat{\mathbf{p}}_t^2)$ *is then*

$$\begin{aligned}
\mathfrak{p}_1^{1,1} &= (1,3,2)', \mathfrak{p}_1^{1,2} = (0,0,0)', \mathfrak{P}_1^{1,1} = (0,3,2)', \mathfrak{P}_1^{1,2} = (0,0,2)', \mathfrak{P}_1^{1,h} = (1/3,1,2/3)';\\
\mathfrak{p}_1^{2,1} &= (0,0,0)', \mathfrak{p}_1^{2,2} = (1,3,2)', \mathfrak{P}_1^{2,1} = (1,0,0)', \mathfrak{P}_1^{2,2} = (1,3,0)', \mathfrak{P}_1^{2,h} = (2/3,2,4/3)';\\
\mathfrak{p}_2^{1,1} &= (2,1,3)', \mathfrak{p}_2^{1,2} = (0,0,0)', \mathfrak{P}_2^{1,1} = (1,1,3)', \mathfrak{P}_2^{1,2} = (0,0,1)', \mathfrak{P}_2^{1,h} = (2/3,1/3,1)';\\
\mathfrak{p}_2^{2,1} &= (0,0,0)', \mathfrak{p}_2^{2,2} = (2,1,3)', \mathfrak{P}_2^{2,1} = (1,0,0)', \mathfrak{P}_2^{2,2} = (2,1,2)', \mathfrak{P}_2^{2,h} = (4/3,2/3,2)'.
\end{aligned}$$

Based on Definitions 1 and 3, we define a *set of feasible personalized prices and quantities*

$$\widehat{S}^A = \{\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M; \widehat{\mathbf{q}}_t\right); t = 1, ..., T\}; \tag{2.1}$$

note that a given set of observations S^A generally enables multiple specifications of $\widehat{S}^A$.

Using the notation $\widehat{S}^A$ we can specify the *generalized axiom of revealed preference* ($GARP$), which we translate towards our specific setting. Varian (1982) introduced the $GARP$ condition for individually rational behavior under observed prices and quantities; i.e. he showed that it is a necessary and sufficient nonparametric condition for maximizing a single non-satiated utility function under a given budget constraint. We focus on the same condition in terms of feasible personalized prices and quantities; we will establish that collective rationality as defined in Definition 2 requires *GARP* consistency for each individual member m.[2]

Definition 4. *Let* $\widehat{S}^A = \{(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M; \widehat{\mathbf{q}}_t); t = 1, ..., T\}$ *be a set of feasible personalized prices and quantities. If* $(\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s \geq (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_t$ *then* $\widehat{\mathbf{q}}_s \ R_0^m \ \widehat{\mathbf{q}}_t$ *('*$\widehat{\mathbf{q}}_s$ *is* directly revealed preferred *to* $\widehat{\mathbf{q}}_t$ *by member m'); and if* $\widehat{\mathbf{q}}_s \ R_0^m \ \widehat{\mathbf{q}}_u$, $\widehat{\mathbf{q}}_u \ R_0^m \ \widehat{\mathbf{q}}_v$, ..., $\widehat{\mathbf{q}}_z \ R_0^m \ \widehat{\mathbf{q}}_t$ *for some (possibly empty) sequence* $(u, v, ..., z)$ *then* $\widehat{\mathbf{q}}_s \ R^m \ \widehat{\mathbf{q}}_t$ *('*$\widehat{\mathbf{q}}_s$ *is* revealed preferred *to* $\widehat{\mathbf{q}}_t$ *by member m'). The set* $\{(\widehat{\mathbf{p}}_t^m; \widehat{\mathbf{q}}_t); t = 1, ..., T\}$ *satisfies* GARP *if* $(\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t \leq (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_s$ *whenever* $\widehat{\mathbf{q}}_s \ R^m \ \widehat{\mathbf{q}}_t$.

Remark that, if the group consists of only a single member ($M = 1$), then $\widehat{S}^A = \{(\mathbf{p}_t; \mathbf{q}_t); t = 1, ..., T\}$ and Definition 4 coincides with the usual *GARP* condition for individually rational behavior. In fact, that *GARP* condition for individually rational behavior can also be interpreted as the nonparametric condition for the unitary household consumption model, which -to recall- treats the household as if it were a single decision maker. This fact that the unitary model can be conceived as a special case of the general collective model (i.e. for $M = 1$) also appears from the next proposition, which provides a nonparametric characterization of collectively rational behavior.

Proposition 1. *Let* $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}), \ t = 1, ..., T\}$ *be a set of observations. The following conditions are equivalent:*

(i) *there exists a combination of* M *concave and continuous utility functions* U^1, ..., U^M *that provide a* collective rationalization *of* S^A;

(ii) *there exists a set of feasible personalized prices and quantities* $\widehat{S}^A$ *such that for each member* $m = 1, ..., M$ *the set* $\{(\widehat{\mathbf{p}}_t^m; \widehat{\mathbf{q}}_t); t = 1, ..., T\}$ *satisfies* GARP;

(iii) *there exists a set of feasible personalized prices and quantities* $\widehat{S}^A$, *numbers* $U_j^m > 0$ *and* $\lambda_j^m > 0$ *such that for all* $s, t \in \{1, ..., T\} : U_s^m - U_t^m \leq \lambda_t^m (\widehat{\mathbf{p}}_t^m)' (\widehat{\mathbf{q}}_s - \widehat{\mathbf{q}}_t)$ *for each member* $m = 1, ..., M$.

Condition (ii) states that collective rationality requires individual rationality (i.e. *GARP* consistency) of each member m in terms of personalized prices and quantities; condition (iii) gives the equivalent 'Afriat inequalities' (see Varian, 1982, for extensive discussion in the context of the unitary model). In general, however, the *true* personalized prices and quantities are unobserved. Therefore, it is only imposed that there must exist at least one set of *feasible* personalized prices and quantities $\widehat{S}^A$ that satisfies the condition. In what follows, we will mainly focus on condition (ii).

Example 3 illustrates the result. This example shows consistency with the condition in Proposition 1 for a data set with two observations. In Section 4 (Example 8), we will give an example with two observations that rejects collective rationality in terms of the condition in Proposition 1 for $M = 2$; this shows that two observations are sufficient for rejecting collective rationality in terms of the condition in Proposition 1. The possibility to reject collective rationality with two observations essentially depends on the available assignable quantity information. Indeed,

[2] Slightly abusing notation, but for ease of exposition, we use $(\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_t = \sum_{l=1}^M (\mathfrak{p}_s^{m,l})' \mathfrak{q}_t^l + \sum_{l=1}^M (\mathfrak{P}_s^{m,l})' \mathfrak{Q}_t^l + (\mathfrak{P}_s^{m,h})' \mathfrak{Q}_t^h$.

Cherchye, De Rock and Vermeulen (2007) show that, if no assignable quantity information is used ($\mathbf{q}_t^{Am} = \mathbf{0}$ for each observation t and each member m), then rejecting collective rationality requires at least three observations (for $M = 2$).

Example 3. *We recapture the situation of Examples 1 and 2, with corresponding observed prices and aggregate quantities. We can verify that this data set satisfies the condition in Proposition 1. For example, consider the set of feasible personalized prices and quantities $\widehat{S}^A$ with $\widehat{\mathbf{q}}_1$ and $\widehat{\mathbf{q}}_2$ specified in Example 1 and $(\widehat{\mathbf{p}}_t^1, \widehat{\mathbf{p}}_t^2)$ specified in Example 2. For these feasible quantities and prices we have that $(\widehat{\mathbf{p}}_1^1)'\widehat{\mathbf{q}}_1(= 13) > (\widehat{\mathbf{p}}_1^1)'\widehat{\mathbf{q}}_2(= 9)$ and $(\widehat{\mathbf{p}}_2^1)'\widehat{\mathbf{q}}_2(= 10) < (\widehat{\mathbf{p}}_2^1)'\widehat{\mathbf{q}}_1(=$ 11), so from Definition 4 it is easily verified that the first member satisfies GARP. Analogously, we find that also the second member satisfies GARP: $(\widehat{\mathbf{p}}_1^2)'\widehat{\mathbf{q}}_1(= 13) < (\widehat{\mathbf{p}}_1^2)'\widehat{\mathbf{q}}_2(= 14)$ and $(\widehat{\mathbf{p}}_2^2)'\widehat{\mathbf{q}}_2(= 16) > (\widehat{\mathbf{p}}_2^2)'\widehat{\mathbf{q}}_1(= 12)$. Since both members satisfy GARP for the given $\widehat{S}^A$, we conclude that the condition in Proposition 1 holds, and thus that there exist utility functions that provide a collective rationalization of this data set.*

2.5. The sharing rule

Importantly in view of our further discussion, the result in Proposition 1 also allows for the following *decentralized* interpretation of collective rationality: collective rationality at the group level (for given S^A) requires individual rationality at the member level (for some $\widehat{S}^A$). Given this, collectively rational consumption behavior can also be represented as the outcome of a two-step allocation procedure: in the first step, the so-called *sharing rule* distributes the aggregate group income across the group members; in the second step, each member optimizes her/his utility subject to the resulting income share and accounting for the member's personalized prices. We remark that this decentralized representation of collectively rational behavior, which follows from the Pareto efficiency assumption regarding the group bargaining process, is formally similar to the well-known decentralization result regarding collective rationality when consumption externalities and public consumption are excluded; see Chiappori (1988, 1992). An important difference of the approach followed in this paper is that each member m's preferences may depend not only on her or his own private consumption, but also on the other members' private consumption as well as public consumption (implying that personalized prices can differ from observed 'market' prices).

In the first step, the sharing rule defines the income shares that are allocated to the different group members. Correspondingly, for a set of feasible personalized prices and quantities $\widehat{S}^A$ that obtains consistency with the collective rationality condition in Proposition 1, we can define *feasible income shares* $\widehat{y}_t^m$ for each member m, which by construction must sum up to the total group budget ($\sum_{m=1}^{M} \widehat{y}_t^m = y_t$).

Definition 5. *Let $\widehat{S}^A = \{(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M; \widehat{\mathbf{q}}_t)\,; t = 1, ..., T\}$ be a set of feasible personalized prices and quantities such that each set $\{(\widehat{\mathbf{p}}_t^m; \widehat{\mathbf{q}}_t)\,;\, t = 1, ..., T\}$, $m = 1, ..., M$, satisfies* GARP. *For $y_t = \mathbf{p}_t'\mathbf{q}_t$ the group income at observation t, this set $\widehat{S}^A$ defines a* feasible income share *for each member m at prices $\mathbf{p}_t$ as $\widehat{y}_t^m = (\widehat{\mathbf{p}}_t^m)'\,\widehat{\mathbf{q}}_t$.*

Example 4 illustrates the definition.

Example 4. *We again consider the data set given in Examples 1 and 2. For the set of feasible personalized prices and quantities* $\widehat{S}^A$ *with* $\widehat{\mathbf{q}}_1$ *and* $\widehat{\mathbf{q}}_2$ *specified in Example 1 and* $(\widehat{\mathbf{p}}_t^1, \widehat{\mathbf{p}}_t^2)$ *specified in Example 2, we obtain* $\widehat{y}_1^1 = (\widehat{\mathbf{p}}_1^1)'\widehat{\mathbf{q}}_1 = 13$, $\widehat{y}_1^2 = (\widehat{\mathbf{p}}_1^2)'\widehat{\mathbf{q}}_1 = 13$ *and* $\widehat{y}_2^1 = (\widehat{\mathbf{p}}_2^1)'\widehat{\mathbf{q}}_2 = 10$, $\widehat{y}_2^2 = (\widehat{\mathbf{p}}_2^2)'\widehat{\mathbf{q}}_2 = 16$. *Observe that* $\widehat{y}_1^1 + \widehat{y}_1^2 = y_1 = 26$ *and* $\widehat{y}_2^1 + \widehat{y}_2^2 = y_2 = 26$.

The second step of the allocation procedure then requires that the quantities $\widehat{\mathbf{q}}_t$ maximize each member m's utility under the budget $\widehat{y}_t^m$ (which, in our set-up, is endogenously defined as $(\widehat{\mathbf{p}}_t^m)'\,\widehat{\mathbf{q}}_t$ for $\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M; \widehat{\mathbf{q}}_t\right)$ in $\widehat{S}^A$). This corresponds to a separate *GARP* condition for each set $\{(\widehat{\mathbf{p}}_t^m; \widehat{\mathbf{q}}_t)\,; t = 1, ..., T\}$.

The sharing rule is a core concept in this two-step representation. It can be interpreted as an indicator for the bargaining power of the individual group members: a higher relative income share of member m $(\widehat{y}_t^m / y_t)$ is then regarded as an indication of increased bargaining power for that member; see Browning, Chiappori and Lewbel (2006). The sharing rule concept is particularly useful in a welfare context, because it is independent of cardinal representations of preferences (in contrast to the bargaining weights μ_t^m in Definition 2). Given this useful interpretation, a main question in what follows concerns the nonparametric recovery of feasible income shares. We will define bounds for the feasible income shares that are independent of the specification of the (data rationalizing) set $\widehat{S}^A$. Intuitively, additional assignable quantity information will generally entail more powerful recovery results. But we will also show that stringent bounds can be obtained even if no assignable quantity information is available.

3. Special cases: testing and recovery

So far, we have considered a collective consumption model that accounts for general utility functions U^m and thus allows for public consumption and externalities of any good. For this case, the necessary and sufficient condition for a collective rationalization in Proposition 1 is difficult to use in practice. More specifically, the member-specific revealed preference relations R_0^m and R^m in Definition 4 are not directly useful since they are nonlinear in the feasible personalized prices $\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M\right)$ and quantities $\widehat{\mathbf{q}}_t$. Given this, we first consider special cases of the general collective consumption model presented in Section 2; they put additional *a priori* structure on member-specific utility functions U^m, which essentially pertains to the nature of the goods in terms of externalities and private/public consumption. In terms of the condition in Proposition 1, for each good they fix *either* the feasible personalized prices *or* the feasible personalized quantities.

We will provide testable necessary and sufficient conditions for collective rationality for these special cases. Starting from these conditions, we can recover the sharing rule, personalized prices and personalized quantities that underlie the observed collective choice behavior. In addition, we can recover, or 'forecast', behavior in new situations. As we will discuss, such testing, recovery and forecasting is possible through mixed integer linear programming (MILP), with binary (or 0-1) variables as the endogenously defined integer variables.[3] As such, practical applications can use efficient solution methods that have been used for formally similar MILP problems; see, for

[3] Closely similar integer programming characterizations have been suggested in the context of Arrovian social welfare functions. See, for example, Sethuraman, Piaw and Vohra (2003).

example, Nemhauser and Wolsey (1999) for a general discussion. Conveniently, efficient MILP solvers have been included in many present-day optimization software packages.

To be precise, nonparametric recovery essentially defines upper and lower *bounds* on the feasible income shares, personalized prices and personalized quantities that hold for all sets $\widehat{S}^A$ providing a collective rationalization of the data; if a specific feasible income share, personalized price or personalized quantity respects these bounds, then there exists a corresponding set $\widehat{S}^A$ that collectively rationalizes the observed set S^A. We illustrate the practical usefulness of the proposed methodology by simple numerical examples. These examples show that precise recovery results (i.e. tight bounds) can be obtained even if there are few observations and no assignable quantity information is available. In practice, of course, we may generally expect more precise recovery when more observations or assignable quantity information can be used.

3.1. Public consumption

In this section, we assume that all private consumption quantities $\mathbf{q}^m$ (without externalities) and $\mathbf{Q}^m$ (with externalities) are zero, which implies $\mathbf{q} = \mathbf{Q}^h$. In terms of the general condition for collective rationality in Definition 2, this means that we consider member-specific utility functions $U^m(\mathbf{q}^m, \mathbf{Q}) = V^m(\mathbf{Q}^h) = V^m(\mathbf{q})$. It is worth emphasizing that this setting is more general than may seem at first sight. *Stricto sensu*, the mere implication is that the (observed) aggregate quantities (*fully*) enter all utility functions; in principle, this allows for private consumption (with externalities) of a particular good e by member m *as long as that good e is exclusively consumed by that member m*. Formally, when using $(\mathbf{z})_e$ as the e-th entry of a vector $\mathbf{z}$, $(\mathbf{Q}^m)_e = (\mathbf{q})_e$ (and thus $(\mathbf{Q}^h)_e = 0$) is empirically equivalent to $(\mathbf{Q}^h)_e = (\mathbf{q})_e$ (and thus $(\mathbf{Q}^m)_e = 0$). This directly relates to our earlier remark that a quantitative but no qualitative distinction can be made between the different components of $\mathbf{Q}$. Further, if externalities are *not* excluded and all goods are fully assignable (i.e. $\mathbf{q} = \sum_{m=1}^{M} \mathbf{Q}^m$ and we observe all quantities $\mathbf{Q}^m$), then an analogous argument obtains that the following method can also be used.

Because we assume that $\mathbf{q}_t = \mathbf{Q}_t^h$ for each observation t, we must focus on sets of feasible personalized prices and quantities $\widehat{S}^A$ with $\mathfrak{Q}_t^h = \mathbf{q}_t$. As a result, the only relevant component of the feasible personalized prices $\widehat{\mathbf{p}}_t^m$ is the vector $\mathfrak{P}_t^{m,h}$, which pertains to the publicly consumed quantities. Given this, the nonparametric necessary and sufficient condition for collective rationality follows directly from Proposition 1.

Corollary 1. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ be a set of observations. For $U^m(\mathbf{q}^m, \mathbf{Q}) = V^m(\mathbf{Q}^h)$, $m = 1, ..., M$, there exists a combination of M concave and continuous utility functions $U^1, ..., U^M$ that provide a* collective rationalization *of S^A if and only if there exist feasible personalized prices $(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M)$ such that for each member m the set $\{(\mathfrak{P}_t^{m,h}; \mathbf{q}_t);\ t = 1, ..., T\}$ satisfies* GARP.

Interestingly, this condition can be reformulated as requiring that the feasible set of a specific MILP problem is non-empty. To see this, we define the binary variables $x_{st}^m \in \{0, 1\}$, with $x_{st}^m = 1$ interpreted as '$\widehat{\mathbf{q}}_s\ R^m\ \widehat{\mathbf{q}}_t$' for a given set of feasible personalized prices and quantities $\widehat{S}^A$. We then have the following result.

Proposition 2. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ be a set of observations. There exist feasible personalized prices $\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M\right)$ such that for each member $m = 1,\ ...,\ M$ the set $\{\left(\mathfrak{P}_t^{m,h}; \mathbf{q}_t\right);\ t = 1, ..., T\}$ satisfies* GARP *if and only if there exist non-negative $\mathfrak{P}_t^{m,h}$, $\widehat{y}_t^m$ and $x_{st}^m \in \{0, 1\}$ that satisfy*

(PP-i) $\mathbf{p}_t = \sum_{m=1}^{M} \mathfrak{P}_t^{m,h}$,

(PP-ii) $\widehat{y}_t^m = \left(\mathfrak{P}_t^{m,h}\right)' \mathbf{q}_t$,

(PP-iii) $\widehat{y}_s^m - \left(\mathfrak{P}_s^{m,h}\right)' \mathbf{q}_t < y_s x_{st}^m$,

(PP-iv) $x_{su}^m + x_{ut}^m \leq 1 + x_{st}^m$, *and*

(PP-v) $\widehat{y}_t^m - \left(\mathfrak{P}_t^{m,h}\right)' \mathbf{q}_s \leq y_t \left(1 - x_{st}^m\right)$.

The interpretation of the different 'personalized price' (PP) constraints is the following. Rule (PP-i) follows from Definition 3 of feasible personalized prices and rule (PP-ii) from Definition 5 of feasible income shares. Rule (PP-iii) implies that, if $\widehat{y}_s^m \geq \left(\mathfrak{P}_s^{m,h}\right)' \mathbf{q}_t$, then we must have $x_{st}^m = 1$ (which corresponds to $\widehat{\mathbf{q}}_s\ R^m\ \widehat{\mathbf{q}}_t$).[4] Rule (PP-iv) imposes transitivity, i.e. $x_{su}^m = 1$ ($\widehat{\mathbf{q}}_s\ R^m\ \widehat{\mathbf{q}}_u$) and $x_{ut}^m = 1$ ($\widehat{\mathbf{q}}_u\ R^m\ \widehat{\mathbf{q}}_t$) imply $x_{st}^m = 1$ ($\widehat{\mathbf{q}}_s\ R^m\ \widehat{\mathbf{q}}_t$). Finally, rule (PP-v) requires that, if $x_{st}^m = 1$ ($\widehat{\mathbf{q}}_s R^m \widehat{\mathbf{q}}_t$), then $\widehat{y}_t^m \leq \left(\mathfrak{P}_t^{m,h}\right)' \mathbf{q}_s$. As such, Proposition 2 defines an operational necessary and sufficient test for collective rationality (under the assumption $U^m\left(\mathbf{q}^m, \mathbf{Q}\right) = V^m\left(\mathbf{Q}^h\right)$): if the MILP constraints (PP-i)-(PP-v) characterize an empty feasible region for the given data set, then a collective rationalization (with only public consumption) of the data is impossible; conversely, if the MILP constraints characterize a non-empty feasible region, then a collective rationalization of the data is certainly possible.

Given this characterization of collective rationality, we can recover upper and lower bounds on feasible income shares and feasible personalized prices that provide a collective rationalization of the set S. To define the upper (or, conversely, lower) bound for the feasible income share of member m, we solve the MILP problem that optimizes the objective $\max \widehat{y}_t^m$ (or $\min \widehat{y}_t^m$) subject to (PP-i)-(PP-v). Similarly, to define the upper (or lower) bound on the feasible personalized price of an individual good e ($1 \leq e \leq n$), we solve the MILP problem that optimizes the objective $\max \left(\mathfrak{P}_t^{m,h}\right)_e$ (or $\min \left(\mathfrak{P}_t^{m,h}\right)_e$) subject to (PP-i)-(PP-v).

Example 5 illustrates the MILP test. It demonstrates that the proposed method can obtain very tight bounds even when the number of observations is small (*in casu* $T = 3$); these tight bounds can be recovered because there is a large variation in the observed prices and aggregate quantities. In the general case, for a given price-quantity variation, we can -of course- expect the bounds to become tighter when more information can be used (e.g. because T gets larger). Such additional information can also include specific hypotheses about the decision structure underlying observed group behavior (*in casu* the sharing rule or feasible personalized prices). In fact, as also shown in Example 5, our approach allows for testing such assumptions.

[4] The strict inequality $\widehat{y}_s^m - \left(\mathfrak{P}_s^{m,h}\right)' \mathbf{q}_t < y_s x_{st}^m$ is difficult to use in MILP. Therefore, in practice we can replace it with $\widehat{y}_s^m - \left(\mathfrak{P}_s^{m,h}\right)' \mathbf{q}_t + \epsilon \leq y_s x_{st}^m$ for ϵ (> 0) arbitrarily small. A similar qualification applies to the constraint (PQ-iv) in Proposition 3.

Example 5. *Consider a two-member household ($M = 2$) that consumes three goods ($n = 3$). Suppose three observations with aggregate quantities and prices (for $0 < \epsilon < 1$)*[5]

$$\begin{aligned}\mathbf{q}_1 &= (1,0,0)', \mathbf{p}_1 = (1+\epsilon, 1, \epsilon/2)', \\ \mathbf{q}_2 &= (0,1,0)', \mathbf{p}_2 = (1, 1+\epsilon, \epsilon/2)', \\ \mathbf{q}_3 &= (0,0,1)', \mathbf{p}_3 = (0.5+\epsilon/2, 0.5+\epsilon/2, 1)'.\end{aligned}$$

As a preliminary step, we note that these prices and quantities imply

$$\begin{aligned}y_1 &= 1+\epsilon,\ \mathbf{p}_1'\mathbf{q}_2 = 1,\ \mathbf{p}_1'\mathbf{q}_3 = \epsilon/2, \\ y_2 &= 1+\epsilon,\ \mathbf{p}_2'\mathbf{q}_1 = 1,\ \mathbf{p}_2'\mathbf{q}_3 = \epsilon/2, \\ y_3 &= 1,\ \mathbf{p}_3'\mathbf{q}_1 = 0.5+\epsilon/2,\ \mathbf{p}_3'\mathbf{q}_2 = 0.5+\epsilon/2.\end{aligned}$$

Step 1. *We first consider the restrictions on the binary variables x^1_{st} and x^2_{st} ($s,t \in \{1,2,3\}$, $s \neq t$) for the current data. As a first result, we must have $x^1_{st} = 1$ or $x^2_{st} = 1$ for any s and t. Specifically, rule (PP-iii) implies*

$$\widehat{y}^1_s - \left(\mathfrak{P}^{1,h}_s\right)' \mathbf{q}_t < y_s x^1_{st} \text{ and } \widehat{y}^2_s - \left(\mathfrak{P}^{2,h}_s\right)' \mathbf{q}_t < y_s x^2_{st}.$$

Combining these two constraints, and using that $\mathbf{p}_s = \mathfrak{P}^{1,h}_s + \mathfrak{P}^{2,h}_s$ (PP-i) and $y_s = \widehat{y}^1_s + \widehat{y}^2_s$ (PP-iii), yields

$$y_s - \mathbf{p}_s'\mathbf{q}_t < 2y_s\left(x^1_{st} + x^2_{st}\right);$$

and thus, because $y_s > \mathbf{p}_s'\mathbf{q}_t$, we necessarily have $x^1_{st} = 1$ or $x^2_{st} = 1$ for any s and t.

As a second result, we obtain that $x^m_{st} = 1$ implies $x^l_{ts} = 1$ ($m,l \in \{1,2\}$, $m \neq l$) for any s and t. Specifically, for $x^m_{st} = 1$ rule (PP-v) entails

$$\widehat{y}^m_t - \left(\mathfrak{P}^{m,h}_t\right)' \mathbf{q}_s \leq 0 \ (= y_t(1 - x^m_{st})).$$

Using $\mathbf{p}_t = \mathfrak{P}^{1,h}_t + \mathfrak{P}^{2,h}_t$, $y_t = \widehat{y}^1_t + \widehat{y}^2_t$ and $y_t > \mathbf{p}_t'\mathbf{q}_s$, this obtains

$$\widehat{y}^l_t > \left(\mathfrak{P}^{l,h}_t\right)' \mathbf{q}_s, \text{ and thus } x^l_{ts} = 1 \text{ because of rule (PP-iii).}$$

As a third result, we cannot have $x^1_{st} = 1$ and *$x^2_{st} = 1$ for any s and t. If $x^1_{st} = 1$ and $x^2_{st} = 1$, then rule (PP-v) requires*

$$\widehat{y}^1_t - \left(\mathfrak{P}^{1,h}_t\right)' \mathbf{q}_s \leq 0 \text{ and } \widehat{y}^2_t - \left(\mathfrak{P}^{2,h}_t\right)' \mathbf{q}_s \leq 0.$$

In turn, using $\mathbf{p}_t = \mathfrak{P}^{1,h}_t + \mathfrak{P}^{2,h}_t$, $y_t = \widehat{y}^1_t + \widehat{y}^2_t$, this yields

$$y_t - \mathbf{p}_t'\mathbf{q}_s \leq 0,$$

[5] To emphasize, we use zero quantities for mathematical elegance. Of course, this use of zero quantities does not affect the core of our arguments in this and following examples.

which is excluded because $y_t > \mathbf{p}_t'\mathbf{q}_s$.

As a fourth result, we cannot have (i) $x_{21}^m = 1$ *and* $x_{31}^l = 1$ *or (ii)* $x_{12}^m = 1$ *and* $x_{32}^l = 1$ $(m \neq l)$. *For example, consider* $x_{21}^m = 1$ *and* $x_{31}^l = 1$. *(The argument for* $x_{12}^m = 1$ *and* $x_{32}^l = 1$ *is directly analogous.) In that case, rule (PP-v) requires*

$$y_1 - \left(\mathfrak{P}_1^{m,h}\right)' \mathbf{q}_2 - \left(\mathfrak{P}_1^{l,h}\right)' \mathbf{q}_3 \leq 0 \; (= \; y_1 \left(2 - x_{21}^m - x_{31}^l\right)),$$

which is excluded because $y_1 > \mathbf{p}_1' \left(\mathbf{q}_2 + \mathbf{q}_3\right)$ *and, by construction,* $\mathbf{p}_1' \left(\mathbf{q}_2 + \mathbf{q}_3\right) \geq \left(\mathfrak{P}_1^{m,h}\right)' \mathbf{q}_2 + \left(\mathfrak{P}_1^{l,h}\right)' \mathbf{q}_3$.

Given these four results, we necessarily obtain $x_{13}^m = x_{12}^m = x_{32}^m = 1$ *and* $x_{23}^l = x_{21}^l = x_{31}^l = 1$. *It is easily verified that this specification satisfies the necessary and sufficient condition in Proposition 2, i.e. the corresponding feasible region defined by rules (PP-i)-(PP-v) is non-empty.*

Step 2. *Next, we consider recovery of the sharing rule. Using rules (PP-ii) and (PP-v) (together with* $\mathbf{p}_3'\mathbf{q}_1 \geq \left(\mathfrak{P}_3^{m,h}\right)' \mathbf{q}_1$ *and* $\mathbf{p}_3'\mathbf{q}_2 \geq \left(\mathfrak{P}_3^{l,h}\right)' \mathbf{q}_2$*, which hold by construction), we obtain*

$$\begin{aligned} x_{13}^m &= 1 \Rightarrow \widehat{y}_3^m \leq \mathbf{p}_3'\mathbf{q}_1 = 0.5 + \epsilon/2 \Rightarrow \widehat{y}_3^l = y_3 - \widehat{y}_3^m \geq 0.5 - \epsilon/2, \\ x_{23}^l &= 1 \Rightarrow \widehat{y}_3^l \leq \mathbf{p}_3'\mathbf{q}_2 = 0.5 + \epsilon/2 \Rightarrow \widehat{y}_3^m = y_3 - \widehat{y}_3^l \geq 0.5 - \epsilon/2; \end{aligned}$$

or, when ϵ *becomes arbitrarily small we obtain very tight bounds (around 0.5) for the feasible income shares* $\widehat{y}_3^1$ *and* $\widehat{y}_3^2$.

Similarly, we get

$$\begin{aligned} x_{32}^m &= 1 \Rightarrow \widehat{y}_2^m \leq \mathbf{p}_2'\mathbf{q}_3 = \epsilon/2 \Rightarrow \widehat{y}_2^l = y_2 - \widehat{y}_2^m \geq 1 - \epsilon/2, \\ x_{31}^l &= 1 \Rightarrow \widehat{y}_1^l \leq \mathbf{p}_1'\mathbf{q}_3 = \epsilon/2 \Rightarrow \widehat{y}_1^m = y_1 - \widehat{y}_1^l \geq 1 - \epsilon/2; \end{aligned}$$

which again obtains tight bounds for $\widehat{y}_t^m$ *and* $\widehat{y}_t^l$ $(t = 1, 2)$ *when* ϵ *gets small. For example,* ϵ *arbitrarily close to zero yields* $\widehat{y}_1^m \approx 1$, $\widehat{y}_1^l \approx 0$ *and* $\widehat{y}_2^m \approx 0$, $\widehat{y}_2^l \approx 1$.

Two remarks are in order. First, this result has a clear interpretation in terms of the 'bargaining power' of the individual members, for which the sharing rule can be interpreted as an indicator. Specifically, consider ϵ *arbitrarily small. In that case, member* m *can be conceived as the (quasi) 'dictator' in situation 1 (i.e. member* m *is solely responsible for the full household budget or* $\widehat{y}_1^m \approx y_1$*) while the other member* l *is the 'dictator' in situation 2 (*$\widehat{y}_2^l \approx y_2$*); in situation 3, finally, the aggregate income is split equally over the two members (*$\widehat{y}_3^1 \approx \widehat{y}_3^2 \approx 0.5y_3$*).*

Second, the proposed method allows for imposing a whole series of additional restrictions on the sharing rule (or, alternatively, for testing specific hypotheses about the sharing rule). Such restrictions preserve the MILP structure as long as they are expressed in linear form. For instance, suppose that in our current example we impose (or assume) that the feasible income share of the wife (member 1) is higher than that of the husband (member 2) in situation 1, i.e. $\widehat{y}_1^1 \geq \widehat{y}_1^2$. *This immediately obtains* $1 - \epsilon/2 \leq \widehat{y}_1^1 \leq 1$, $0 \leq \widehat{y}_1^2 \leq \epsilon/2$ *and* $0 \leq \widehat{y}_2^1 \leq \epsilon/2$, $1 - \epsilon/2 \leq \widehat{y}_2^2 \leq 1$*; and, thus, for* ϵ *arbitrarily small the mere restriction* $\widehat{y}_1^1 \geq \widehat{y}_1^2$ *implies that the wife is the 'dictator' in situation 1 (*$\widehat{y}_1^1 \approx y_1$*) and the husband is the 'dictator' in situation 2 (*$\widehat{y}_2^2 \approx y_2$*).*

Alternatively, one can put upper and lower bounds (or test corresponding assumptions) on the relative income share of some member m in situation t, i.e. $\underline{y}_t^m \leq \widehat{y}_t^m / y_t \leq \overline{y}_t^m$ for $\underline{y}_t^m, \overline{y}_t^m \in [0,1]$; the linear nature of these constraints is consistent with the MILP formulation given above. For instance, our result implies that any lower bound $\underline{y}_t^m > \epsilon/2$ for some m and all t will be rejected for this specific data structure. Finally, additional sharing rule restrictions can impose a specific relationship between feasible income shares of the same member m in different situations (e.g. time periods). For instance, suppose that we assume in the current example that the feasible income share of the wife must be higher in situation 1 than in situation 2, i.e. $\widehat{y}_1^1 \geq \widehat{y}_2^1$; this directly obtains $1-\epsilon/2 \leq \widehat{y}_1^1 \leq 1$, $0 \leq \widehat{y}_2^1 \leq \epsilon/2$ and $0 \leq \widehat{y}_1^2 \leq \epsilon/2$, $1-\epsilon/2 \leq \widehat{y}_2^2 \leq 1$.

Step 3. *Let us then consider recovery of the feasible personalized prices. As a starting point, we use our conclusion for the feasible income shares, which can be summarized as*

$$\begin{aligned} 1-\epsilon/2 &\leq \widehat{y}_1^m \leq 1 \text{ and } 0 \leq \widehat{y}_1^l \leq \epsilon/2, \\ 0 &\leq \widehat{y}_2^m \leq \epsilon/2 \text{ and } 1-\epsilon/2 \leq \widehat{y}_2^l \leq 1, \\ 0.5-\epsilon/2 &\leq \widehat{y}_3^1 \leq 0.5+\epsilon/2 \text{ and } 0.5-\epsilon/2 \leq \widehat{y}_3^2 \leq 0.5+\epsilon/2. \end{aligned}$$

For the given data structure, this implies (using (PP-ii))

$$\begin{aligned} 1-\epsilon/2 &\leq \left(\mathfrak{P}_1^{m,h}\right)_1 \leq 1 \text{ and } 0 \leq \left(\mathfrak{P}_1^{l,h}\right)_1 \leq \epsilon/2, \\ 0 &\leq \left(\mathfrak{P}_2^{l,h}\right)_2 \leq \epsilon/2 \text{ and } 1-\epsilon/2 \leq \left(\mathfrak{P}_2^{m,h}\right)_2 \leq 1, \\ 0.5-\epsilon/2 &\leq \left(\mathfrak{P}_3^{1,h}\right)_3 \leq 0.5+\epsilon/2 \text{ and } 0.5-\epsilon/2 \leq \left(\mathfrak{P}_3^{2,h}\right)_3 \leq 0.5+\epsilon/2. \end{aligned}$$

We thus get very tight bounds for $\left(\mathfrak{P}_t^{m,h}\right)_t$ and $\left(\mathfrak{P}_t^{l,h}\right)_t$ when ϵ gets arbitrarily small. To illustrate the impact of additional structure, suppose that $\left(\mathfrak{P}_1^{1,h}\right)_1 > \left(\mathfrak{P}_1^{2,h}\right)_1$, i.e. the wife contributes more to the first good in situation 1. For ϵ arbitrarily small, this mere restriction implies that the wife 'pays' (quasi) everything of the first good in situation 1 ($\left(\mathfrak{P}_1^{1,h}\right)_1 \approx (\mathbf{p}_1)_1$), while the husband pays everything of the second good in situation 2 ($\left(\mathfrak{P}_2^{2,h}\right)_2 \approx (\mathbf{p}_2)_2$); finally, in situation 3 the expenditure for the third good is equally split ($\left(\mathfrak{P}_3^{1,h}\right)_3 \approx \left(\mathfrak{P}_3^{2,h}\right)_3 \approx 0.5\,(\mathbf{p}_3)_3$).

3.2. Private consumption without externalities

In this section, we consider the specific case that excludes externalities and public consumption ($\mathbf{Q}_t = (\mathbf{0}, ..., \mathbf{0}, \mathbf{0})$); i.e. all goods are private and group members are of the 'egoistic' type. In terms of the general condition for collective rationality in Definition 2, this means that we consider member-specific utility functions $U^m(\mathbf{q}^m, \mathbf{Q}) = V^m(\mathbf{q}^m)$. At this point, it is worth noting that this case actually also encompasses a wider class of member-specific utilities that model 'altruism' in a specific way: it also includes so-called 'caring preferences', which correspond to utility functions $U^m(\mathbf{q}^m, \mathbf{Q}) = W^m\left(V^1(\mathbf{q}^1), ..., V^M\left(\mathbf{q}^M\right)\right)$ that depend not only on member m's own 'egoistic' utility but also on the other member l's utility defined in terms of $\mathbf{q}^l$. Chiappori (1992) argues that every Pareto efficient outcome in terms of caring preferences

(W^m) is also Pareto efficient in terms of egoistic preferences (V^m). In other words, under Pareto efficiency the empirical implications of caring preferences are indistinguishable from those of egoistic preferences.

As a preliminary note, we recall that under the stated conditions, which imply $\mathfrak{Q}_t = (\mathbf{0}, ..., \mathbf{0}, \mathbf{0})$ for the feasible personalized quantities, any set $\widehat{S}^A$ of feasible personalized prices and quantities must meet

$$\mathbf{q}_t = \sum_{m=1}^{M} \mathfrak{q}_t^m \text{ with } \mathfrak{q}_t^m \geq \mathbf{q}_t^{Am}. \tag{3.1}$$

This implies the 'trivial' bounds

$$(\mathbf{0} \leq)\ \mathbf{q}_t^{Am} \leq \mathfrak{q}_t^m \leq \left(\mathbf{q}_t - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_t^{Al}\right)\ (\leq \mathbf{q}_t). \tag{3.2}$$

We will show that collective rationality imposes additional restrictions on the personalized private quantities that can imply (substantially) tighter bounds than those in (3.2). We will also demonstrate that very tight bounds can be obtained even if no assignable quantity information is available.

Like before, we first formulate the necessary and sufficient condition for collective rationality that is relevant in the present case. This condition follows directly from Proposition 1.

Corollary 2. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ be a set of observations. For $U^m(\mathbf{q}^m, \mathbf{Q}) = V^m(\mathbf{q}^m)$, $m = 1$, ..., M, there exists a combination of M concave and continuous utility functions U^1, ..., U^M that provide a* collective rationalization *of S^A if and only if there exist feasible personalized quantities $\widehat{\mathbf{q}}_t$ with $\mathbf{q}_t = \sum_{m=1}^{M} \mathfrak{q}_t^m$ such that for each member m the set $\{(\mathbf{p}_t; \mathfrak{q}_t^m); t = 1, ..., T\}$ satisfies* GARP.

Once more, we can reformulate this condition as requiring that the feasible set of a specific MILP problem is non-empty. This is contained in the following result.

Proposition 3. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ be a set of observations. There exist feasible personalized quantities $\widehat{\mathbf{q}}_t$ with $\mathbf{q}_t = \sum_{m=1}^{M} \mathfrak{q}_t^m$ such that for each member $m = 1$, ..., M the set $\{(\mathbf{p}_t; \mathfrak{q}_t^m); t = 1, ..., T\}$ satisfies* GARP *if and only if there exist non-negative $\mathfrak{q}_t^m$, $\widehat{y}_t^m$ and $x_{st}^m \in \{0, 1\}$ that satisfy*

(PQ-i) $\mathbf{q}_t = \sum_{m=1}^{M} \mathfrak{q}_t^m$,

(PQ-ii) $\mathbf{q}_t^{Am} \leq \mathfrak{q}_t^m$,

(PQ-iii) $\widehat{y}_t^m = \mathbf{p}_t' \mathfrak{q}_t^m$,

(PQ-iv) $\widehat{y}_s^m - \mathbf{p}_s' \mathfrak{q}_t^m < y_s x_{st}^m$,

(PQ-v) $x_{su}^m + x_{ut}^m \leq 1 + x_{st}^m$, *and*

(PQ-vi) $\widehat{y}_t^m - \mathbf{p}_t' \mathfrak{q}_s^m \leq y_t (1 - x_{st}^m)$.

The different 'personalized quantity' (PQ) constraints have a similar interpretation as the personalized price (PP) constraints in Proposition 2. Rules (PQ-i) and (PQ-ii) repeat the constraints in (3.1). Rule (PQ-iii) follows from Definition 5 of feasible income shares. Rule

(PQ-iv) implies that, if $\widehat{y}_t^m \geq \mathbf{p}_s' \mathfrak{q}_t^m$, then we must have $x_{st}^m = 1$ (which corresponds to $\widehat{\mathbf{q}}_s \ R^m \ \widehat{\mathbf{q}}_t$). Rule (PQ-v) imposes transitivity. Finally, rule (PP-vi) requires that, if $x_{st}^m = 1$ ($\widehat{\mathbf{q}}_s R^m \widehat{\mathbf{q}}_t$), then $\widehat{y}_t^m \leq \left(\mathfrak{P}_t^{m,h}\right)' \mathbf{q}_s$. As such, Proposition 3 defines a necessary and sufficient condition for collective rationality (under the assumption $U^m(\mathbf{q}^m, \mathbf{Q}) = V^m(\mathbf{q}^m)$) that can be tested through MILP. Given this characterization, we can define upper and lower bounds on feasible income shares and feasible personalized quantities by solving MILP optimization problems. For example, an upper (or, conversely, lower) bound on the feasible personalized quantity of an individual good e ($1 \leq e \leq n$) is obtained by optimizing the objective $\max(\mathfrak{q}_t^m)_e$ (or $\min(\mathfrak{q}_t^m)_e$) subject to (PQ-i)-(PQ-vi).

Example 6 illustrates the result. It demonstrates that the proposed method can obtain very tight bounds when the number of observations is small. In addition, it shows that such tight bounds can be obtained for the feasible personalized quantities even if no assignable quantity information is available. Analogously to before, additional information can include specific hypotheses regarding the group decision process (e.g. the sharing rule and the assignable quantities). Again, our approach effectively allows for testing such assumptions.

Example 6. *We recapture the situation of Example 5, with corresponding observed prices and aggregate quantities. This example does not include assignable quantity information, so that (PQ-ii) does not add information.*

As for the feasible income shares, an analogous reasoning as in Steps 1 and 2 of Example 5 yields the conclusion (for $m \neq l$)

$$\begin{aligned}
1 - \epsilon/2 &\leq \widehat{y}_1^m \leq 1 \text{ and } 0 \leq \widehat{y}_1^l \leq \epsilon/2, \\
0 &\leq \widehat{y}_2^m \leq \epsilon/2 \text{ and } 1 - \epsilon/2 \leq \widehat{y}_2^l \leq 1, \\
0.5 - \epsilon/2 &\leq \widehat{y}_3^1 \leq 0.5 + \epsilon/2 \text{ and } 0.5 - \epsilon/2 \leq \widehat{y}_3^2 \leq 0.5 + \epsilon/2.
\end{aligned}$$

Focusing on the feasible personalized quantities, this implies (using (PQ-iii))

$$\begin{aligned}
1 - \epsilon/2 &\leq (1+\epsilon)(\mathfrak{q}_1^m)_1 \leq 1 \text{ and } 0 \leq (1+\epsilon)\left(\mathfrak{q}_1^l\right)_1 \leq \epsilon/2, \\
0 &\leq (1+\epsilon)\left(\mathfrak{q}_2^l\right)_2 \leq \epsilon/2 \text{ and } 1 - \epsilon/2 \leq (1+\epsilon)(\mathfrak{q}_2^m)_2 \leq 1, \\
0.5 - \epsilon/2 &\leq \left(\mathfrak{q}_3^1\right)_3 \leq 0.5 + \epsilon/2 \text{ and } 0.5 - \epsilon/2 \leq \left(\mathfrak{q}_3^2\right)_3 \leq 0.5 + \epsilon/2.
\end{aligned}$$

We thus obtain

$$\begin{aligned}
(1 - \epsilon/2)/(1+\epsilon) &\leq (\mathfrak{q}_1^m)_1 \leq 1/(1+\epsilon) \text{ and } 0 \leq \left(\mathfrak{q}_1^l\right)_1 \leq \epsilon/\left(2(1+\epsilon)\right), \\
0 &\leq \left(\mathfrak{q}_2^l\right)_2 \leq \epsilon/\left(2(1+\epsilon)\right) \text{ and } (1-\epsilon/2)/(1+\epsilon) \leq (\mathfrak{q}_2^m)_2 \leq 1/(1+\epsilon), \\
0.5 - \epsilon/2 &\leq \left(\mathfrak{q}_3^1\right)_3 \leq 0.5 + \epsilon/2 \text{ and } 0.5 - \epsilon/2 \leq \left(\mathfrak{q}_3^2\right)_3 \leq 0.5 + \epsilon/2.
\end{aligned}$$

This yields very tight bounds for $(\mathfrak{q}_t^m)_t$ and $\left(\mathfrak{q}_t^l\right)_t$ when ϵ gets arbitrarily small. To illustrate the impact of additional structure, suppose that $(\mathfrak{q}_1^1)_1 > (\mathfrak{q}_1^2)_1$, i.e. the wife consumes more of the first good in situation 1. For ϵ arbitrarily small, this sole restriction immediately obtains that the wife consumes (quasi) everything of the first good in situation 1 ($(\mathfrak{q}_1^1)_1 \approx (\mathbf{q}_1)_1$), while the husband consumes everything of the second good in situation 2 ($(\mathfrak{q}_2^2)_2 \approx (\mathbf{q}_2)_2$); finally, in situation 3 the third good is equally split ($(\mathfrak{q}_3^1)_3 \approx (\mathfrak{q}_3^2)_3 \approx 0.5\,(\mathbf{q}_3)_3$).

3.3. Extensions

To conclude, we indicate that the previous methodology allows for a number of interesting extensions. As illustrated in our examples, it enables us to test specific hypotheses regarding the collective decision process (e.g. on the sharing rule, in Example 5). In what follows, and without being exhaustive, we point out three additional applications. Because we believe the formal analogy with the previous discussion is fairly easy, we restrict to sketching the main arguments.

1. The previous discussion on testing and recovery restricted to (i) the case without private consumption (with recovery of feasible personalized prices) and (ii) the case without externalities and public consumption (with recovery of feasible personalized quantities). In practice, intermediate cases may also be considered. In this respect, it is worth emphasizing that arguments directly analogous to those given above apply to the case in which we can *a priori* identify each good as either exclusively publicly consumed or exclusively privately consumed (without externalities). (Importantly, if we want to obtain similar MILP formulations as before, no good can be partly privately consumed (without externalities) *and* partly publicly consumed.)

   Formally, such an intermediate case implies $\mathbf{Q}^m = \mathbf{0}$ for all m and $\left(\sum_{m=1}^{M} \mathbf{q}^m\right) \odot \mathbf{Q}^h = \mathbf{0}$ (with $\odot$ the Hadamard or 'element-by-element' product). Thus, we must consider sets $\widehat{S}^A$ of feasible personalized prices and quantities with $\mathfrak{Q}_t^m = \mathbf{0}$ and $\left(\sum_{m=1}^{M} \mathfrak{q}_t^m\right) \odot \mathfrak{Q}_t^h = \mathbf{0}$. Given this, an analogous argumentation as before establishes a necessary and sufficient condition for collective rationality that is essentially a 'hybrid' version of the conditions in Propositions 2 and 3. This characterization obtains an operational MILP test for collective rationality, which in turn enables the recovery of bounds on feasible income shares, feasible personalized prices (for the goods that are publicly consumed) and feasible personalized quantities (for the goods that are privately consumed). Finally, data structures similar to the one in Example 5 can demonstrate the potential of the method (i) to obtain very tight bounds even if no assignable quantity information is available and the number of observations is small, and (ii) to test alternative hypotheses (regarding the sharing rule, feasible personalized prices and feasible personalized quantities).

2. Another interesting application concerns the recovery, or 'forecasting', of individual members' consumption (and thus also aggregate group consumption) in a new situation defined in terms of new prices $\mathbf{p}_N$ and group income y_N. For example, this allows us to compare individual consumption and the bargaining power of some member m (captured by $\widehat{y}_t^m / y_t$ for each situation t) in alternative (observed and new) situations. To introduce this application, we recapture the case that excludes externalities and public consumption, which corresponds to $U^m(\mathbf{q}^m, \mathbf{Q}) = V^m(\mathbf{q}^m)$; this allows us to develop similar MILP formulations as before.

   Essentially, we have to construct the set of feasible personalized quantities $\mathfrak{q}_N^m$ that meet $y_N = \mathbf{p}_N' \mathbf{q}_N$ for aggregate group demand $\mathbf{q}_N = \sum_{m=1}^{M} \mathfrak{q}_N^m$, such that each set $\{(\mathbf{p}_t; \mathfrak{q}_t^m);\ t = 1, ..., T\} \cup \{(\mathbf{p}_N; \mathfrak{q}_N^m)\}$ satisfies *GARP*; every feasible specification of the $\mathfrak{q}_N^m$ defines feasible income shares $\widehat{y}_N^m = \mathbf{p}_N' \mathfrak{q}_N^m$. The MILP characterization of the set of feasible personalized quantities $\mathfrak{q}_N^m \in \mathbb{R}_+^n$ and feasible income shares $\widehat{y}_N^m \in \mathbb{R}_+$ is then a direct extension of the one in Proposition 3. We remark that this MILP formulation easily includes alternative prior assumptions regarding the consumption quantities in the new situation. For example, we can restrict the proportions of the group quantities that will be consumed by the individual group members (e.g. in a two-member household, such a restriction can impose that a particular good is exclusively consumed by the husband); this boils down to imposing additional constraints of the form $\underline{\alpha}^m \odot \mathbf{q}_N \leq \mathfrak{q}_N^m \leq \overline{\alpha}^m \odot \mathbf{q}_N$ for $\underline{\alpha}^m$, $\overline{\alpha}^m \in \mathbb{R}_+^n$. Generally, such additional restrictions preserve the MILP structure as long as they are expressed in linear form.

3. A final extension involves recovering the behavior of some member m in a new situation defined in terms of new prices $\mathbf{p}_N$ and a given utility level (i.e. the same utility level for member m as the initially observed bundle $\mathbf{q}_I$, $I \in \{1, ..., T\}$). In such an application, the recovered bounds on the feasible income shares can subsequently be used for constructing member-specific cost-of-living indices (corresponding to the prices $\mathbf{p}_N$ and the same utility level as $\mathbf{q}_I$ for m); compare with Varian (1982), Blow and Crawford (2001) and Blundell, Browning and Crawford (2003), who conduct nonparametric cost-of-living analyses in a unitary setting. For example, this may be useful for comparing the cost-of-living of individuals living alone with that of the same individuals living in a multi-member household; compare with Browning, Chiappori and Lewbel (2006), who address such a question by using parametric methods.

   Again, we consider the situation without externalities or public consumption. In this case, we need to characterize the feasible personalized quantities $\mathfrak{q}_N^m$ (which define feasible income shares $\widehat{y}_N^m = \mathbf{p}_N' \mathfrak{q}_N^m$ for member m) that simultaneously meet the following conditions: (i) the set $\{(\mathbf{p}_t; \mathfrak{q}_t^m)\,;\, t = 1, ..., T\} \cup \{(\mathbf{p}_N; \mathfrak{q}_N^m)\}$ must satisfy *GARP*; (ii) for member m the quantities $\mathfrak{q}_N^m$ must be 'equally good' as those corresponding to the initial bundle $\mathbf{q}_I$. Once more, the MILP formulation of condition (i) follows immediately from Proposition 3. As for condition (ii), we must include the additional restrictions $x_{sI}^m \leq x_{sN}^m$ and $x_{It}^m \leq x_{Nt}^m$: the first constraint implies $x_{sN}^m = 1$ if $x_{sI}^m = 1$ (which complies with '$\widehat{\mathbf{q}}_s$ $R^m \widehat{\mathbf{q}}_N$ if $\widehat{\mathbf{q}}_s$ $R^m \widehat{\mathbf{q}}_I$'), and the second constraint requires $x_{Nt}^m = 1$ if $x_{It}^m = 1$ ('$\widehat{\mathbf{q}}_N$ $R^m \widehat{\mathbf{q}}_t$ if $\widehat{\mathbf{q}}_I$ $R^m \widehat{\mathbf{q}}_t$'). Using this characterization, we can define bounds on the feasible income shares ($\widehat{y}_N^m$) and the feasible personalized quantities ($(\mathfrak{q}_N^m)_e$ for some good e) by solving MILP problems. Like before, we can also include alternative prior assumptions regarding the consumption quantities in $\mathfrak{q}_N^m$; such additional restrictions preserve the MILP structure as long as they are expressed in linear form.

4. The general case: testing

We next turn to the general collective consumption model defined in Definition 2, which accounts for public consumption and externalities of any good. For this case, we do not observe the 'true' specification of *either* the feasible personalized prices $\left(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M\right)$ *or* the feasible personalized quantities $\widehat{\mathbf{q}}_t$. This is in sharp contrast with the special cases in Section 3; for the condition in Proposition 1, it excludes developing an equivalent MILP formulation similar to the ones in Propositions 2 and 3. The observed set S^A usually allows for multiple specifications of both

$(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M)$ and $\widehat{\mathbf{q}}_t$, and each such specification implies different restrictions in terms of the relations R_0^m and R^m.

Therefore, in the following we focus on testable restrictions on R_0^m and R^m that are directly expressed in terms of observed prices and quantities, and that do not refer to a specific $\widehat{S}^A$. This obtains a testable necessary condition for data consistency with the general collective model which solely uses the prices and quantities that are effectively observed; this condition extends the necessary condition of Cherchye, De Rock and Vermeulen (2007) by accounting for assignable quantity information. Consistent with our (MILP-based) approach in Section 3, we show that this necessary condition can be reformulated in integer programming (IP) terms, which is again attractive for practical applications. In addition, as we will discuss in Section 5, it provides a useful basis for sharing rule recovery in the case of the general collective consumption model.

4.1. A preliminary result

Before introducing our testable necessary condition, we present a lemma that provides the starting point of our approach. It implies that we can start from the set S^A for specifying restrictions on R_0^m. Moreover, the equivalence results imply that we cannot do better when using only the set of observations S^A (rather than a specific $\widehat{S}^A$).

Lemma 1. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}),\ t = 1, ..., T\}$ be a set of observations. For any s and t, we have the following two equivalence results (for $m \in \{1, ..., M\}$):*

(i) $\left[\text{for all sets } \widehat{S}^A\text{, there exists } m : \widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t\right] \Leftrightarrow [\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\mathbf{q}_t]$;

(ii) *for m:* $\left[\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t \text{ for all sets } \widehat{S}^A\right] \Leftrightarrow \left[\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right)\right]$.

Rule (i) does not use assignable quantity information and pertains to the Pareto efficient nature of group behavior in the collective model; for $M = 2$, it equals Lemma 1 of Cherchye, De Rock and Vermeulen (2007). Specifically, if the group has chosen $\mathbf{q}_s$ when $\mathbf{q}_t$ was equally available ($\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\mathbf{q}_t$), then we always have that, *independently* of the specification of the set $\widehat{S}^A$, at least one group member m must prefer the former (personalized) quantities to the latter (i.e. $\widehat{\mathbf{q}}_s\ R_0^m\ \widehat{\mathbf{q}}_t$); the identity of member m depends on the specification of $\widehat{S}^A$ that is used.

Rule (ii) does use assignable quantity information, and shows that this effectively allows us to 'assign' preference relations to an individual group member m; such 'assignable' relations for member m hold for *any* specification of $\widehat{S}^A$. It uses that, by construction, $(\widehat{\mathbf{p}}_s^m)'\,\widehat{\mathbf{q}}_s \geq \mathbf{p}_s'\mathbf{q}_s^{Am}$ and $\mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right) \geq (\widehat{\mathbf{p}}_s^m)'\,\widehat{\mathbf{q}}_t$, so that

$$\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right) \text{ implies } (\widehat{\mathbf{p}}_s^m)'\,\widehat{\mathbf{q}}_s \geq (\widehat{\mathbf{p}}_s^m)'\,\widehat{\mathbf{q}}_t \text{ for any } \widehat{S}^A;$$

and it follows from this last inequality that $\widehat{\mathbf{q}}_s\ R_0^m\ \widehat{\mathbf{q}}_t$ for any $\widehat{S}^A$. In words, the 'minimal' (assignable) expenditures of member m in observation s exceed the 'maximal' expenditures of that member for bundle t (under the prices $\mathbf{p}_s$), which implies that member m 'reveals' her/his preference for bundle s over bundle t. In the limiting case that all goods are fully assignable ($\mathbf{q}_t = \sum_{m=1}^M \mathbf{q}_t^{Am}$, which implies $\mathbf{q}_t^m = \mathbf{q}_t^{Am}$) the right hand side of rule (ii) reduces to $\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\mathbf{q}_t^{Am}$ and, thus, all member-specific preference relations are assignable.

Rule (ii) implies that assignable quantity information implies additional empirical restrictions as compared to the limiting case with $\mathbf{q}_t^{Am} = \mathbf{0}$ for each m and t (i.e. no assignable quantity information, so that only rule (i) can be used). To illustrate, we consider Example 7, which shows that assignable quantity information allows for recovering preference relations R_0^m even if rule (i) is not applicable. This suggests that, in general, the use of assignable quantity information can obtain a testable condition for collective rationality that is stronger than the one that is solely based on rule (i), which was originally presented by Cherchye, De Rock and Vermeulen (2007). Consequently, such assignable information can entail a more powerful empirical analysis. The more stringent testable condition will be discussed next.

Example 7. *Consider a two-member household ($M = 2$) with a wife (member 1) and a husband (member 2) that consumes three goods ($n = 3$). Suppose two observations with aggregate quantities and prices*

$$\begin{aligned} \mathbf{q}_1 &= (4,2,2)', \mathbf{p}_1 = (4,5,1)', \\ \mathbf{q}_2 &= (2,4,2)', \mathbf{p}_2 = (1,4,5)', \end{aligned}$$

and assignable quantities

$$\begin{aligned} \mathbf{q}_1^{A1} &= (3,0,0)' \text{ and } \mathbf{q}_1^{A2} = (0,1,2)'; \\ \mathbf{q}_2^{A1} &= (0,1,1)' \text{ and } \mathbf{q}_2^{A2} = (1,3,0)'. \end{aligned}$$

We then obtain $\mathbf{p}_1'\mathbf{q}_1^{A1}(= 12) > \mathbf{p}_1'\left(\mathbf{q}_2 - \mathbf{q}_2^{A2}\right)(= 11)$ and thus, on the basis of rule (ii) in Lemma 1, we can conclude $\widehat{\mathbf{q}}_1 \ R_0^1 \ \widehat{\mathbf{q}}_2$ for every set $\widehat{S}^A$; i.e., for every *feasible specification of the personalized prices and quantities '$\widehat{\mathbf{q}}_1$ is directly revealed preference to $\widehat{\mathbf{q}}_2$' by the wife. On the other hand, we have $\mathbf{p}_1'\mathbf{q}_1(= 28) < \mathbf{p}_1'\mathbf{q}_2(= 30)$, and so we cannot use rule (i) (to conclude $\widehat{\mathbf{q}}_1$ $R_0^m \ \widehat{\mathbf{q}}_2$, $m = 1$ or 2, for any set $\widehat{S}^A$).*

4.2. Testable necessary condition

The basic idea of our testable condition is to formulate restrictions on 'feasible' specifications of the relations R_0^m and R^m, which are expressed in terms of the observed information summarized in the set S^A. Such feasible specifications are then referred to as *hypothetical* relations. Specifically, we specify $\mathbf{q}_s \ H_0^m \ \mathbf{q}_t$ if we *hypothesize* $\widehat{\mathbf{q}}_s \ R_0^m \ \widehat{\mathbf{q}}_t$; H^m denotes the transitive closure of the relation H_0^m. We say that a collective rationalization of the data in the sense of Definition 2 is impossible if there does not exist a feasible specification of these hypothetical relations that satisfies the restrictions defined in the following Proposition 4. This defines a necessary condition for collectively rational behavior as characterized in Proposition 1. In addition, as we will discuss, it implies an operational test for data consistency with the general collective consumption model.[6]

[6] We note that rules (iv) and (v) refine rules (iii) and (iv) originally defined by Cherchye, De Rock and Vermeulen (2007, Proposition S3); the new rules (iv) and (v) strengthen the original rules (iii) and (iv) in that they imply them as a special case.

Proposition 4. *Suppose that there exists a combination of M utility functions U^1, ..., U^M that provide a* collective rationalization *of the set of observations $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}), t = 1, ..., T\}$. Then there exist* hypothetical *relations H_0^m and H^m, $m = 1, ..., M$, such that:*
(i) *if $\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\mathbf{q}_t$, then $\mathbf{q}_s\ H_0^m\ \mathbf{q}_t$ for some m;*
(ii) *if $\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^{M} \mathbf{q}_t^{Al}\right)$, then $\mathbf{q}_s\ H_0^m\ \mathbf{q}_t$;*
(iii) *if $\mathbf{q}_s\ H_0^m\ \mathbf{q}_u$, $\mathbf{q}_u\ H_0^m\ \mathbf{q}_v$, ..., $\mathbf{q}_z\ H_0^m\ \mathbf{q}_t$ for some (possibly empty) sequence $(u, v, ..., z)$, then $\mathbf{q}_s\ H^m\ \mathbf{q}_t$;*
(iv) *for $M^* \leq M$ and $\mathbf{M} \subsetneq \{1, ..., M\}$: if $\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\left(\sum_{k=1}^{M^*} \mathbf{q}_{t_k}\right)$ and for all $m \in \mathbf{M}$ we have $\mathbf{q}_{t_{k(m)}}\ H^m\ \mathbf{q}_s$ for some $k(m) \leq M^*$, then $\mathbf{q}_s\ H_0^l\ \mathbf{q}_{t_k}$ for some $l \notin \mathbf{M}$ and $k \leq M^*$;*
(v) *for $M^* \leq M$ and $\mathbf{M} \subsetneq \{1, ..., M\}$: if $\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\left(\sum_{k=1}^{M^*} \mathbf{q}_{t_k}\right)$ and for all $m \in \mathbf{M}$ we have $\mathbf{q}_{t_{k(m)}}\ H^m\ \mathbf{q}_s$ for some $k(m) \leq M^*-1$, then $\mathbf{q}_s\ H_0^l\ \mathbf{q}_{t_{M^*}}$ for some $l \notin \mathbf{M}$;*
(vi) *for $M^* \leq M$: if for all m we have $\mathbf{q}_{s_{k(m)}}\ H^m\ \mathbf{q}_t$ for some $k(m) \leq M^*$, then $\mathbf{p}_t'\mathbf{q}_t \leq \sum_{k=1}^{M^*} \mathbf{p}_t'\mathbf{q}_{s_k}$;*
(vii) *if $\mathbf{q}_s\ H^m\ \mathbf{q}_t$, then $\mathbf{p}_t'\mathbf{q}_t^{Am} \leq \mathbf{p}_t'\left(\mathbf{q}_s - \sum_{l=1,l\neq m}^{M} \mathbf{q}_s^{Al}\right)$.*

Hence, a collective rationalization of the set S^A requires that there exists at least one specification of the relations H_0^m and H^m consistent with rules (i)-(vii) in this proposition. Of course, in general there may be multiple feasible specifications of H_0^m and H^m that obtain consistency with rules (i)-(vii).

Let us then interpret the different rules in Proposition 4. First, rules (i) and (ii) follow immediately from Lemma 1, when replacing the relations R_0^m by their hypothetical counterparts H_0^m. Next, rule (iii) defines the transitive closures H^m of the relations H_0^m.

The interpretation of the remaining rules (iv) to (vii) pertains to the very nature of the collective model, which -to recall- explicitly recognizes the multi-person nature of the group decision process. Rules (iv) and (v) compare $\mathbf{q}_s$ to (combinations of) M^* *different* bundles $\mathbf{q}_{t_k}$. First, rule (iv) expresses that, if all members $m \in \mathbf{M}$ prefer some $\mathbf{q}_{t_{k(m)}}$ over $\mathbf{q}_s$ for the (sum) bundle $\sum_{k=1}^{M^*} \mathbf{q}_{t_k}$ not more expensive than $\mathbf{q}_s$, then the choice of $\mathbf{q}_s$ can be rationalized only if another member $l \notin \mathbf{M}$ prefers $\mathbf{q}_s$ over some $\mathbf{q}_{t_k}$. Next, rule (v) expresses that, if (aggregate) $\mathbf{q}_s$ is more expensive than the (sum) bundle $\sum_{k=1}^{M^*} \mathbf{q}_{t_k}$, while each member $m \in \mathbf{M}$ prefers $\mathbf{q}_{t_{k(m)}}$ (with $k(m) \leq M^*-1$) over $\mathbf{q}_s$, then the only possibility for rationalizing the choice of $\mathbf{q}_s$ is that another member $l \notin \mathbf{M}$ prefers $\mathbf{q}_s$ over the remaining bundle $\mathbf{q}_{t_{M^*}}$.

Rules (i) to (v) define restrictions on the relations H_0^m and H^m. For a specification of these relations, rules (vi) and (vii) define the corresponding upper cost bound conditions. Rule (vi) complements rules (iv) and (v). It states that, if each member m prefers $\mathbf{q}_{s_{k(m)}}$ ($k(m) \leq M^*$) over $\mathbf{q}_t$, then the choice of $\mathbf{q}_t$ can be rationalized only if it is not more expensive than the (sum) bundle $\sum_{k=1}^{M^*} \mathbf{q}_{s_k}$. In this expression, we can have $M^* < M$ because it is possible that $s_{k(m)} = s_{k(l)}$ for $l \neq m$ (i.e. 'members m and l both prefer the same bundle $\mathbf{q}_{s_{k(m)}}$ ($= \mathbf{q}_{s_{k(l)}}$) over $\mathbf{q}_t$').

Finally, rule (vii) reveals that the assignable quantity information makes it possible to formulate *separate* upper cost bound conditions for the individual group members, whereas the upper cost bound defined by rule (vi) corresponds to all members simultaneously. More specifically, rule (vii) expresses that, if member m prefers $\mathbf{q}_s$ over $\mathbf{q}_t$, then the latter choice can be rationalized only if $\mathbf{p}_t'\mathbf{q}_t^{Am} \leq \mathbf{p}_t'\left(\mathbf{q}_s - \sum_{l=1,l\neq m}^{M} \mathbf{q}_s^{Al}\right)$.

An important observation is that the condition in Proposition 4 is empirically rejectable, i.e. we can find data that do not satisfy the condition. Cherchye, De Rock and Vermeulen (2007) provide such an example for the case when there is no assignable quantity information ($\mathbf{q}_t^{Am} = \mathbf{0}$ for each m and t). They show that, in such a situation, $M+1$ observations and $M+1$ goods are both necessary and sufficient to reject the general collective consumption model for M-member group behavior.

Example 8 provides an illustration that uses assignable quantity information; it obtains rejection of collective rationality with two observations ($T = 2$) for a two-member household ($M = 2$). In fact, when assignable quantity information is available, then in general it suffices to have two observations and goods to reject collective rationality of M-member group behavior. Specifically, it is easy to verify that in the limiting case that all goods are fully assignable (i.e. $\mathbf{q}_t = \sum_{m=1}^{M} \mathbf{q}_t^{Am}$ and thus $\mathbf{q}_t^m = \mathbf{q}_t^{Am}$), the condition in Proposition 4 boils down to *GARP* consistency of the sets $\{(\mathbf{p}_t; \mathbf{q}_t^{Am}); t = 1, ..., T\}$ corresponding to each member m; and this is rejected if $\mathbf{p}_t'\mathbf{q}_t^{Am} > \mathbf{p}_t'\mathbf{q}_s^{Am}$ and $\mathbf{p}_s'\mathbf{q}_s^{Am} > \mathbf{p}_s'\mathbf{q}_t^{Am}$ (for some m and t, s), which requires only two goods and two observations. This again illustrates that, in general, more assignable quantity information will entail more powerful results.

Example 8. *We recapture the situation of a household with two members and three goods in Example 7. This specific data structure implies $\mathbf{q}_1 \; H_0^1 \; \mathbf{q}_2$ because of rule (ii). But then rule (vii) is not met because $\mathbf{p}_2'\mathbf{q}_2^{A1}(= 9) > \mathbf{p}_2'\left(\mathbf{q}_1 - \mathbf{q}_1^{A2}\right)(= 8)$. Thus, we conclude a violation of the condition in Proposition 4.*

4.3. Integer programming formulation

In this section, we show that the necessary condition in Proposition 4 can be reformulated in IP terms, which is attractive from an operational point of view. As a preliminary step, we define the binary variables $a_{M^*}[s; t_1, ..., t_{M^*}]$, $b^m[s,t] \in \{0,1\}$ (for $m, M^* \in \{1, ..., M\}$, $s, t \in \{1, ..., T\}$) and $\{t_1, ..., t_{M^*}\} \subseteq \{1, ..., T\}$:

$a_{M^*}[s; t_1, ..., t_{M^*}] = 1$ if $\mathbf{p}_s'\mathbf{q}_s \geq \sum_{k=1}^{M^*} \mathbf{p}_s'\mathbf{q}_{t_k}$ and $a_{M^*}[s; t_1, ..., t_{M^*}] = 0$ otherwise;

$b^m[s;t] = 1$ if $\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^{M} \mathbf{q}_t^{Al}\right)$ and $b^m[s;t] = 0$ otherwise.

These variables capture the available information that is used in rules (i)-(vii) of Proposition 4. We next formulate these rules as IP constraints. To do so, we again define the binary variables $x_{st}^m \in \{0,1\}$. As we focus on the necessary condition, which is expressed in hypothetical preference relations, $x_{st}^m = 1$ must now be interpreted as '$\mathbf{q}_s \; H^m \; \mathbf{q}_t$'.

Given this, rules (i) and (ii) are equivalent to, respectively,

(IP-i) $\sum_{m=1}^{M} x_{st}^m \geq a_1[s;t]$ and **(IP-ii)** $x_{st}^m \geq b^m[s;t]$.

The constraint (IP-i) implies that, if $a_1[s;t] = 1$, then $x_{st}^m = 1$ for some m. Similarly, the constraint (IP-ii) implies that, if $b^m[s;t] = 1$, then $x_{st}^m = 1$.

Rule (iii) corresponds to

(IP-iii) $x^m_{su} + x^m_{ut} \leq 1 + x^m_{st}$.

Thus, if $x^m_{su} = x^m_{ut} = 1$ then $x^m_{st} = 1$, which imposes transitivity.

To provide the IP formulation of rule (iv), for each combination $(t_1, ..., t_{M^*})$ and any subset $\mathbf{M} \subsetneq \{1, ..., M\}$, we consider $k(m) \leq M^*$ for all $m \in \mathbf{M}$. Given this, rule (iv) complies with

(IP-iv) for each $(t_1, ..., t_{M^*})$, $\mathbf{M} \subsetneq \{1, ..., M\}$ and correspondingly defined $k(m) \leq M^*$ for all $m \in \mathbf{M}$:

$$a_{M^*}[s; t_1, ..., t_{M^*}] + \sum_{m \in \mathbf{M}} x^m_{t_{k(m)}s} \leq |\mathbf{M}| + \sum_{l \notin \mathbf{M}} \sum_{k=1}^{M^*} x^l_{st_k}.$$

This imposes that, if $a_{M^*}[s; t_1, ..., t_{M^*}] = 1$ and for all $m \in \mathbf{M}$ we have $x^m_{t_{k(m)}s} = 1$, then we must have $x^l_{st_k} = 1$ for some $l \notin \mathbf{M}$ and $k \leq M^*$.

Similarly, for rule (v) we consider $k(m) \leq M^*{-}1$ for all $m \in \mathbf{M}$. Rule (v) then corresponds to

(IP-v) for each $(t_1, ..., t_{M^*})$, $\mathbf{M} \subsetneq \{1, ..., M\}$ and correspondingly defined $k(m) \leq M^*{-}1$ for all $m \in \mathbf{M}$:

$$a_{M^*}[s; t_1, ..., t_{M^*}] + \sum_{m \in \mathbf{M}} x^m_{t_{k(m)}s} \leq |\mathbf{M}| + \sum_{l \notin \mathbf{M}} x^l_{st_{M^*}}.$$

The interpretation is that, if $a_{M^*}[s; t_1, ..., t_{M^*}] = 1$ and for all $m \in \mathbf{M}$ we have $x^m_{t_{k(m)}s} = 1$, then for the remaining t_{M^*} we must have $x^l_{st_{M^*}} = 1$ for some $l \notin \mathbf{M}$.

Next, to define the IP formulation of rule (vi), for each M^* we consider every combination $(s_{k(1)}, ..., s_{k(M)})$ with $k(m) \leq M^*$ for all $m \in \mathbf{M}$; note that we can have $k(m) = k(l)$ $(m \neq l)$ so that $M^* \leq M$. Given this, rule (vi) is equivalent to

(IP-vi) for each $(s_{k(1)}, ..., s_{k(M)})$ with $k(m) \leq M^*$ for all m :

$$\sum_{m=1}^{M} x^m_{s_{k(m)}t} \leq M - a_{M^*}[t; s_1, ..., s_{M^*}] \text{ if } \mathbf{p}'_t \mathbf{q}_t < \sum_{k=1}^{M^*} \mathbf{p}'_t \mathbf{q}_{t_k}.$$

This gives the upper cost bound condition that applies to all members simultaneously.[7] For every possible combination $\left(s_{k(1)}, ..., s_{k(M)}\right)$, it specifies that, if for each m we have $x^m_{s_{k(m)}t} = 1$, then it must be that $a_{M^*}[t; s_1, ..., s_{M^*}] = 0$ (using that $k(m) \leq M^*$ for all m).

Finally, rule (vii) complies with

(IP-vii) $x^m_{st} \leq 1 - b^m[t; s]$ if $\mathbf{p}'_t \mathbf{q}^{Am}_t > \mathbf{p}'_t \left(\mathbf{q}_s - \sum_{l=1, l \neq m}^{M} \mathbf{q}^{Al}_s \right)$.

This imposes that, if $x^m_{st} = 1$, then it must hold that $b^m[t; s] = 0$. It specifies an upper cost bound condition for each individual member m.

As such, testing consistency with collective rationality requires checking whether the constraints (IP-i)-(IP-vii) characterize a non-empty feasible region (for $x^m_{st} \in \{0, 1\}$). Every feasible specification of the binary variables x^m_{st} corresponds to a specification of the relations H^m consistent with rules (i)-(vii) in Proposition 4.

To conclude, we provide the simple numerical Example 9 as an illustration.

---

[7] For completeness, we note that the inequality constraint in (IP-vi) can equivalently be formulated as $\sum_{m=1}^{M} x^m_{s_{k(m)}t} \leq M - a^s_{M^*}[t; s_1, ..., s_{M^*}]$ for $a^s_{M^*}[t; s_1, ..., s_{M^*}] = 1$ if $\mathbf{p}'_t \mathbf{q}_t > \sum_{k=1}^{M^*} \mathbf{p}'_t \mathbf{q}_{t_k}$ and $a^s_{M^*}[t; s_1, ..., s_{M^*}] = 0$ otherwise. An analogous qualification applies to (IP-vii).

Example 9. *We recapture the situation of a household with two members and three goods in Examples 7 and 8. For this specific data structure the inequality* $\mathbf{p}_1'\mathbf{q}_1^{A1}(=12) > \mathbf{p}_1'\left(\mathbf{q}_2 - \mathbf{q}_2^{A2}\right)$ $(=11)$ *implies* $b^1[1;2] = 1$*, and thus* $x_{12}^1 = 1$ *because of (IP-ii). Given this, (IP-vii) requires* $b^1[2;1] = 0$*, which contradicts* $\mathbf{p}_2'\mathbf{q}_2^{A1}(=9) > \mathbf{p}_2'\left(\mathbf{q}_1 - \mathbf{q}_1^{A2}\right)(=8)$*. As such, we obtain an empty feasible region and conclude that the condition in Proposition 4 is violated.*

4.4. Remarks

As mentioned before, the necessary condition developed in this section extends the necessary condition of Cherchye, De Rock and Vermeulen (2007) by including the possible use of assignable quantity information. This directly obtains that the necessary condition is in general not sufficient. The result follows from Example 2 of Cherchye, De Rock and Vermeulen (2007): for $M = 2$, it presents data (without assignable quantity information) that satisfy the condition but cannot be collectively rationalized in the sense of Proposition 1. For general T, the condition is sufficient if all goods are fully assignable ($\mathbf{q}_t = \sum_{m=1}^{M} \mathbf{q}_t^{Am}$, which implies $\mathbf{q}_t^m = \mathbf{q}_t^{Am}$). In addition, it is sufficient if $M = 1$, when it reduces to the usual GARP condition for individually rational behavior (see Varian, 1982). (In fact, the same qualification holds for the conditions in Propositions 2 and 3.) In these cases, the feasible personalized prices and quantities are all observed.

We see at least the following arguments to motivate our focus on the necessary condition in Proposition 4. First, the condition is always sufficient for T not 'too large', which depends on the number of members M and the assignable quantity information. For example, Cherchye, De Rock and Vermeulen (2007) argue that, for $M = 2$, it is sufficient if $T \leq 4$ and if no assignable quantity information is available ($\mathbf{q}_t^{Am} = \mathbf{0}$ for each t and m); but, *ceteris paribus*, the condition is no longer sufficient for $T = 5$. Similar results can be derived for $M > 2$ and $\mathbf{q}_t^{Am} \neq \mathbf{0}$ for some t and m.

Next, Cherchye, De Rock and Vermeulen (2007, Proposition S4) present a testable sufficient condition for the general collective model, which can easily be adapted to the current set-up with assignable quantity information. (Alternative sufficient conditions for the general model are those in Propositions 2 and 3, which are necessary and sufficient for special cases of the general model.) Interestingly, these authors also provide a 'convergence' argument for the case without assignable quantity information, which states that the empirical implications of their sufficient condition generally come closer to those of the necessary condition in Proposition 4 when the sample size gets larger; i.e. both conditions become equally powerful for larger T. Again, this convergence argument is easily adapted to account for assignable quantity information.

Another argument relates to the subtle difference between 'collectively rational household behavior' and 'data consistency with collectively rational household behavior'. While inconsistency with a necessary condition *necessarily* implies collectively irrational behavior, consistency with a sufficient condition in general does not imply collectively rational behavior; it *only* implies data consistency with collectively rational consumption behavior. In other words, any sufficient condition only allows for 'non-rejection' (but not for 'acceptance') of the collective rationality hypothesis.

Finally, our focus on the necessary condition falls in line with the very nature of the nonparametric approach that we follow, which typically focuses on the minimal (or 'necessity') empirical restrictions that can be obtained from the available data. In Section 5, we will argue that the necessary condition provides a powerful basis for recovering bounds on feasible income shares that underlie the observed (collectively rational) group behavior. Specifically, we derive bounds that must be respected by *any set $\widehat{S}^A$ that provides a collective rationalization of the data* in the sense of Proposition 1; this is a direct consequence of the fact that our recovery method starts from a necessary condition (and not a sufficient condition) for collective rationality.

5. The general case: recovery

In this section, we will show that the necessary condition in Proposition 4 allows for sharing rule recovery in the case of the general collective consumption model, and that such recovery is possible through MILP. As a preliminary remark, we indicate that our use of a necessary condition for collective rationality as a starting point entails a subtle difference with the recovery results in Section 3, which were based on necessary and sufficient conditions for (special cases of) collective rationality. In particular, a specific feasible income share for member m that respects the upper and lower bounds that we will characterize must no longer necessarily correspond to a set $\widehat{S}^A$ that collectively rationalizes the data. Such necessary correspondence is only the case when the necessary condition is also sufficient; see also our discussion at the end of Section 4. Still, we believe that the bounds that we will present do provide useful information, even when the necessary condition is not sufficient: in our opinion, the fact that they necessarily bound each feasible income share for member m that is defined by a data rationalizing set $\widehat{S}^A$ is the more important property in view of practical applications.

5.1. An independence result

Before addressing recovery of the sharing rule, we argue that the necessary condition in Proposition 4 provides a useful basis for recovery. To do so, we show that, if assignable quantity information can be used, this necessary condition for $M = M'$ with $M' > 1$ is independent of the *GARP* condition for the unitary model, which -to recall- coincides with the collective rationality condition for $M = 1$ (i.e. individual rationality): data consistency with the necessary condition for M' members is neither necessary nor sufficient for data consistency with the unitary *GARP* condition. Because Varian (1982) used this unitary *GARP* condition for addressing recovery questions in the unitary setting, we believe this convincingly motivates our use of the necessary condition as a basis for (*in casu* sharing rule) recovery in the case of the general collective consumption model.

Example 10 illustrates the independence result by presenting (i) data that do not satisfy the individual rationality condition for $M = 1$ while they do pass the necessary condition for $M = 2$ and (ii) data that satisfy the individual rationality condition for $M = 1$ but not the necessary condition for $M = 2$; similar examples can be conceived for any $M' > 1$. As such, considering additional members should not necessarily imply a weaker test; the necessary condition for multi-member collective rationality should not have weaker empirical implications than the *GARP* condition for individual rationality if there is assignable quantity information. (In this respect, it is also worth recalling that the *GARP* condition for individual rationality coincides with the necessary condition in Proposition 4 for $M = 1$.)

At this point, we note that Chiappori (1988; Examples 1-2 on p. 76-77) obtains a similar independence conclusion for his collective labor supply model in the case of egoistic agents.

The approach followed in this paper (including Example 10) clarifies that the independence essentially relates to the assignable quantity information that is available (in Chiappori's Examples 1-2, this information pertains to observed leisure of the individual household members in combination with the assumption of egoistic preferences). In other words, if no assignable quantity information is available, then we do have that data consistency with the necessary condition for M' members is *always* necessary for data consistency with the *GARP* condition for 1 member; but this is no longer the case if we can assign private consumption quantities (without externalities) to individual household members.

Example 10. *As a first illustration, we recapture the data (observed prices and aggregate quantities) in Examples 1 and 2. On the one hand, Example 3 concludes that these data satisfy the necessary and sufficient condition in Proposition 1 (and thus also the necessary condition in Proposition 4) when using the corresponding assignable quantities for $M = 2$; in words, a data rationalization in terms of 'two-member rationality' is possible. On the other hand, it is easily verified that these data do not satisfy the unitary* GARP *condition (i.e. there is a single decision maker and, thus, $\widehat{S}^A = \{(\mathbf{p}_t; \mathbf{q}_t)\,;\, t = 1, 2\}$). Specifically, we have that $\mathbf{p}_1'\mathbf{q}_1(= 26) > \mathbf{p}_1'\mathbf{q}_2(= 23)$ while $\mathbf{p}_2'\mathbf{q}_1(= 23) < \mathbf{p}_2'\mathbf{q}_2(= 26)$, which obtains the result. This shows that, in general, consistency with the necessary condition for M' ($M' > 1$) members does not necessarily imply consistency with the unitary* GARP *condition.*

As a second illustration, we recapture the data (observed prices and aggregate quantities) in Example 7. On the one hand, Examples 8 and 9 conclude that these data do not satisfy the necessary condition in Proposition 4 when using the corresponding assignable quantities for $M = 2$; in words, a data rationalization in terms of 'two-member rationality' is impossible. On the other hand, it is easily verified that these data do satisfy the unitary GARP *condition. Specifically, we have that $\mathbf{p}_1'\mathbf{q}_1(= 28) < \mathbf{p}_1'\mathbf{q}_2(= 30)$ and $\mathbf{p}_2'\mathbf{q}_2(= 28) > \mathbf{p}_2'\mathbf{q}_1(= 22)$, which gives the result. More generally, this shows that, with assignable quantity information, consistency with the unitary* GARP *condition does not necessarily imply consistency with the necessary condition for $M = M'$ with $M' > 1$.*

Given this indepence and the fact that Varian (1982) used the unitary *GARP* condition for unitary recovery, we conclude that the necessary condition in Proposition 4 provides a powerful basis for nonparametrically addressing sharing rule recovery. An important result in our following discussion is that this condition can obtain precise recovery *even* if *no* assignable quantity information is available. Still, at this point we want to stress that we do -of course-expect assignable quantity information to yield important value-added in practical applications. Specifically, the foregoing discussion makes clear that additional assignable quantity information generally yields a more stringent necessary condition, which in turn obtains more precise recovery results. For brevity, we will not explicitly illustrate recovery with assignable quantity information in what follows, but we believe the analogy with the example that will be given (without assignable quantity information) is fairly straightforward.

5.2. Sharing rule recovery

Essentially, we define *bounds* on the feasible income shares $\widehat{y}_t^m$ defined in Definition 5 in terms of the feasibility restrictions implied by the necessary condition in Proposition 4. As a preliminary step, we recall the definitional fact

$$\sum\nolimits_{m=1}^{M} \widehat{y}_t^m = y_t, \tag{5.1}$$

which holds for any specification of the set of feasible personalized prices and quantities $\widehat{S}^A = \{(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M; \widehat{\mathbf{q}}_t); t = 1, ..., T\}$. Combination of (5.1) with $\mathbf{q}_t^l \geq \mathbf{q}_t^{Al}$ for all l defines 'trivial' upper and lower bounds on $\widehat{y}_t^m$ as

$$(0 \leq)\ \mathbf{p}_t' \mathbf{q}_t^{Am} \leq \widehat{y}_t^m \leq \mathbf{p}_t' \left(\mathbf{q}_t - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_t^{Al} \right)\ \ (\leq y_t); \tag{5.2}$$

and, of course, these initial bounds will generally be tighter when more assignable quantity information is available. In the following, we will show that collective rationality (summarized in terms of the necessary condition in Proposition 4) implies additional restrictions on the income shares $\widehat{y}_t^m$ that, in general, can imply (substantially) tighter bounds than those defined in (5.2).

To sketch the basic idea, we first consider a *specific* set $\widehat{S}^A$ and specify the restrictions on feasible income shares $\widehat{y}_t^m$ that are implied by the corresponding relations R^m (without explicitly considering the corresponding specification of prices $\widehat{\mathbf{p}}_t^m$ and quantities $\widehat{\mathbf{q}}_t^m$ in $\widehat{S}^A$). For the given set $\widehat{S}^A$, Definition 4 requires $\widehat{y}_t^m = (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t \leq (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_s$ whenever $\widehat{\mathbf{q}}_s\ R^m\ \widehat{\mathbf{q}}_t$. Using $(\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_s \leq \mathbf{p}_t' \left(\mathbf{q}_s - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_s^{Al} \right)$, we obtain that

$$\widehat{y}_t^m \leq \mathbf{p}_t' \left(\mathbf{q}_s - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_s^{Al} \right) \text{ whenever } \widehat{\mathbf{q}}_s R^m \widehat{\mathbf{q}}_t.$$

As such, if $\mathbf{p}_t' \left(\mathbf{q}_s - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_s^{Al} \right) < \mathbf{p}_t' \left(\mathbf{q}_t - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_t^{Al} \right)$, this obtains an upper bound on the income share $\widehat{y}_t^m$ of member m that is lower than the trivial upper bound $\mathbf{p}_t' \left(\mathbf{q}_t - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_t^{Al} \right)$ in (5.2). Next, similarly constructed upper bounds for $\widehat{y}_t^l$ $(l \neq m)$ define a lower bound for $\widehat{y}_t^m$ that can be higher than the trivial lower bound $\mathbf{p}_t' \mathbf{q}_t^{Am}$ in (5.2).

In practice, we do not observe a specific $\widehat{S}^A$ and corresponding relations R^m; but the approach developed in Section 4 allows for defining restrictions on 'feasible' specifications of the R^m, which we defined in terms of the hypothetical relations H^m (Proposition 4). Similar to before, we avoid using a specific $\widehat{S}^A$. That is, we replace the relations R^m by their hypothetical counterparts H^m in the above argument and, consequently, consider specifications of the hypothetical relations H^m that are consistent with the rules (i)-(vii) in Proposition 4. Starting from our earlier IP formulation of the necessary condition, we reformulate the hypothetical relations H^m in terms of the binary variables $x_{st}^m \in \{0, 1\}$ (with, to recall, $x_{st}^m = 1$ interpreted as '$\mathbf{q}_s\ H^m\ \mathbf{q}_t$'). This obtains the following result.

Proposition 5. *Let $S^A = \{(\mathbf{p}_t, \mathbf{q}_t; \mathbf{q}_t^{A1}, ..., \mathbf{q}_t^{AM}), t = 1, ..., T\}$ be a set of observations. For any set $\widehat{S}^A$ that satisfies condition (ii) in Proposition 1, the corresponding feasible income shares $\widehat{y}_t^m$, $m = 1, ..., M$, meet*

(SR-i) $\sum_{m=1}^{M} \widehat{y}_t^m = y_t,$

(SR-ii) $\mathbf{p}_t' \mathbf{q}_t^{Am} \leq \widehat{y}_t^m,$ *and*

(SR-iii) $\widehat{y}_t^m - \mathbf{p}_t' \left(\mathbf{q}_s - \sum\nolimits_{l=1, l \neq m}^{M} \mathbf{q}_s^{Al} \right) \leq y_t \left(1 - x_{st}^m \right),$

for $x^m_{st} \in \{0,1\}$ consistent with (IP-i)-(IP-vii).

In this result, the fact that we consider $x^m_{st} \in \{0,1\}$ consistent with (IP-i)-(IP-vii) implies that we focus on feasible income shares defined by hypothetical relations H^m that meet rules (i)-(vii) in Proposition 4. Given this, the interpretation of the 'sharing rule' (SR) constraints is as follows. First, the constraint (SR-i) imposes (5.1) while the constraint (SR-ii) implies the trivial bounds defined in (5.2); they are -of course- linear in nature. Finally, the constraint (SR-iii) imposes $\widehat{y}^m_t \leq \mathbf{p}'_t \left(\mathbf{q}_s - \sum_{l=1,l\neq m}^{M} \mathbf{q}^{Al}_s\right)$ if $x^m_{st} = 1$ (which corresponds to $\mathbf{q}_s H^m \mathbf{q}_t$). Similar to before, given this characterization of the set of feasible income shares, one can define upper (or lower) bounds on the income share for each member m by solving the MILP problem that optimizes the objective $\max \widehat{y}^m_t$ (or $\min \widehat{y}^m_t$) subject to (IP-i)-(IP-vii) and (SR-i)-(SR-iii).

Example 11 illustrates the result by recapturing the data structure of Examples 5 and 6, which considered special cases of the collective model. Interestingly, even though we impose minimal a priori structure (in terms of preferences and assignable quantity information), we get exactly the same sharing rule bounds as in these special cases. This also shows that the proposed method can yield very tight bounds even if no assignable quantity information can be used and the sample is small. Of course, we can generally expect the bounds to become tighter when more information can be used (including additional assignable quantity information and/or more observations). Like before, such additional information can also involve additional restrictions (or testable assumptions) on the sharing rule; compare with our discussion of Example 5.

Example 11. *We recapture the situation of Example 5, with corresponding observed prices and aggregate quantities. Since there is no assignable quantity information ($\mathbf{q}^{Am}_t = \mathbf{0}$ for $m = 1,2$ and $t = 1,2,3$), the trivial bounds in (SR-ii) merely imply $0 \leq \widehat{y}^m_t \leq y_t$ for each m and t and, thus, (SR-ii) is redundant in view of (SR-i). As a preliminary step, recall that these prices and quantities imply*

$$\begin{aligned}
y_1 &= 1+\epsilon,\ \mathbf{p}'_1\mathbf{q}_2 = 1,\ \mathbf{p}'_1\mathbf{q}_3 = \epsilon/2, \\
y_2 &= 1+\epsilon,\ \mathbf{p}'_2\mathbf{q}_1 = 1,\ \mathbf{p}'_2\mathbf{q}_3 = \epsilon/2, \\
y_3 &= 1,\ \mathbf{p}'_3\mathbf{q}_1 = 0.5+\epsilon/2,\ \mathbf{p}'_3\mathbf{q}_2 = 0.5+\epsilon/2.
\end{aligned}$$

On the one hand, because $\mathbf{p}'_s\mathbf{q}_s > \mathbf{p}'_s\mathbf{q}_t$ and $\mathbf{p}'_t\mathbf{q}_t > \mathbf{p}'_t\mathbf{q}_s$ for all $s,t = 1,2,3$ we must have $\mathbf{q}_s H^m \mathbf{q}_t$ and $\mathbf{q}_t H^l \mathbf{q}_s$ ($t \neq s$ and $m \neq l$); this follows from rules (i) and (iv) in Proposition 4. On the other hand, rule (vi) in Proposition 4 implies that we cannot have $\mathbf{q}_2 H^m \mathbf{q}_1$ and $\mathbf{q}_3 H^l \mathbf{q}_1$ ($m \neq l$) because $\mathbf{p}'_1\mathbf{q}_1 > \mathbf{p}'_1(\mathbf{q}_2 + \mathbf{q}_3)$; and, similarly, because $\mathbf{p}'_2\mathbf{q}_2 > \mathbf{p}'_2(\mathbf{q}_1 + \mathbf{q}_3)$ we cannot have $\mathbf{q}_1 H^m \mathbf{q}_2$ and $\mathbf{q}_3 H^l \mathbf{q}_2$. Summarizing, we must always have $\mathbf{q}_1 H^m \mathbf{q}_3$, $\mathbf{q}_1 H^m \mathbf{q}_2$, $\mathbf{q}_3 H^m \mathbf{q}_2$ and $\mathbf{q}_2 H^l \mathbf{q}_3$, $\mathbf{q}_2 H^l \mathbf{q}_1$, $\mathbf{q}_3 H^l \mathbf{q}_1$. Or, using the IP formulation, we necessarily obtain $x^m_{13} = x^m_{12} = x^m_{32} = 1$ and $x^l_{23} = x^l_{21} = x^l_{31} = 1$. It is easily verified (e.g. using the IP formulation) that this specification satisfies the necessary condition in Proposition 4 (and we recall that this condition is also sufficient for $T = 3$ if there is no assignable quantity information).

Using (SR-i) and (SR-iii), this specification of the hypothetical relations implies

$$\begin{aligned}
x^m_{13} &= 1 \Rightarrow \widehat{y}^m_3 \leq \mathbf{p}'_3\mathbf{q}_1 = 0.5+\epsilon/2 \Rightarrow \widehat{y}^l_3 = y_3 - \widehat{y}^m_3 \geq 0.5-\epsilon/2, \\
x^l_{23} &= 1 \Rightarrow \widehat{y}^l_3 \leq \mathbf{p}'_3\mathbf{q}_2 = 0.5+\epsilon/2 \Rightarrow \widehat{y}^m_3 = y_3 - \widehat{y}^l_3 \geq 0.5-\epsilon/2.
\end{aligned}$$

Similarly, we get

$$\begin{aligned} x_{32}^{m} &= 1 \Rightarrow \widehat{y}_{2}^{m} \leq \mathbf{p}_{2}'\mathbf{q}_{3} = \epsilon/2 \Rightarrow \widehat{y}_{2}^{l} = y_{2} - \widehat{y}_{2}^{m} \geq 1 - \epsilon/2, \\ x_{31}^{l} &= 1 \Rightarrow \widehat{y}_{1}^{l} \leq \mathbf{p}_{1}'\mathbf{q}_{3} = \epsilon/2 \Rightarrow \widehat{y}_{1}^{m} = y_{1} - \widehat{y}_{1}^{l} \geq 1 - \epsilon/2. \end{aligned}$$

This obtains tight bounds for $\widehat{y}_t^m$ and $\widehat{y}_t^l$ ($t = 1, 2, 3$) when ϵ gets small. For example, ϵ arbitrarily close to zero yields $\widehat{y}_1^m \approx 1$, $\widehat{y}_1^l \approx 0$ and $\widehat{y}_2^m \approx 0$, $\widehat{y}_2^l \approx 1$, while $\widehat{y}_3^1 \approx \widehat{y}_3^2 \approx 0.5$.

As a concluding remark, we indicate that the proposed method is not readily adapted for recovering the feasible personalized prices and quantities for the general collective consumption model under consideration. In fact, such non-recoverability applies for some good as soon as we cannot identify it *a priori* either as exclusively privately consumed without externalities, or as exclusively publicly consumed. (The other cases have been covered in Section 3.) To illustrate, we consider such a good e that is privately consumed and characterized by externalities; we exclude public consumption and private consumption without externalities to keep the argument simple. For this good, *neither* the feasible personalized quantities $((\mathfrak{Q}_t^m)_e$ for each member $m)$ *nor* the feasible personalized prices $((\mathfrak{P}_t^{l,m})_e$ for each pair of members m and $l)$ are fixed *a priori*. By using the sharing rule bounds, which -to recall- can still be recovered in this case, the method subsequently allows for bounding the product $(\mathfrak{P}_t^{l,m})_e (\mathfrak{Q}_t^m)_e$ of the feasible personalized prices and quantities. But it does *not* allow for separately bounding the constituent factors $(\mathfrak{P}_t^{l,m})_e$ and $(\mathfrak{Q}_t^m)_e$. Example 12 illustrates the argument.

Example 12. *We recapture the situation of Example 11, with corresponding observed prices and aggregate quantities. Suppose that the good 3 is privately consumed and characterized by externalities, i.e. $(\mathfrak{q}_t^1)_3 = (\mathfrak{q}_t^2)_3 = \left(\mathfrak{Q}_t^h\right)_3 = 0$ for $t = 1, 2, 3$. The conclusion of Example 11 then implies*

$$\begin{aligned} 0.5 - \epsilon/2 &\leq \left(\mathfrak{P}_3^{1,1}\right)_3 \left(\mathfrak{Q}_3^1\right)_3 + \left(\mathfrak{P}_3^{1,2}\right)_3 \left(\mathfrak{Q}_3^2\right)_3 \leq 0.5 + \epsilon/2 \text{ and} \\ 0.5 - \epsilon/2 &\leq \left(\mathfrak{P}_3^{2,1}\right)_3 \left(\mathfrak{Q}_3^1\right)_3 + \left(\mathfrak{P}_3^{2,2}\right)_3 \left(\mathfrak{Q}_3^2\right)_3 \leq 0.5 + \epsilon/2; \end{aligned}$$

and, clearly, these non-linear constraints do not impose separate bounds for $\left(\mathfrak{P}_3^{1,m}\right)_3$, $\left(\mathfrak{P}_3^{2,m}\right)_3$ and $(\mathfrak{Q}_3^m)_3$. [Evidently, directly similar arguments can be constructed for the goods 1 and 2.]

Interestingly, this limitation of our method complies with a similar conclusion in the parametric literature (see Chiappori and Ekeland, 2005, for a detailed discussion). In that literature, existing results fail to obtain 'identifiability' (of the decision structure underlying the observed collective consumption behavior) in exactly the same cases in which our method fails to recover (separate bounds on) feasible personalized prices and quantities.

6. Summary and concluding remarks

We have extended the nonparametric 'revealed preference' methodology for analyzing collective consumption behavior, so that it can be used for empirically addressing welfare-related questions that are specific to the collective model. First, we established a nonparametric characterization of collectively rational behavior that includes the possibility that assignable quantity information is available. Starting from this characterization, we have next presented nonparametric testing and recovery tools for special cases of the collective model, which impose specific structure in terms of consumption externalities and public consumption, as well as for the general collective model, which imposes minimal *a priori* structure. We have shown that testing and recovery is possible through integer programming (IP and MILP) with binary (or 0-1) variables as the integer variables. Finally, while we have argued that additional assignable quantity information generally yields more powerful recovery results, our examples also demonstrate that the proposed methodology can obtain precise recovery even if no assignable quantity information is available.

Given all this, the next crucial step consists of bringing the proposed methodology to real-life consumption data, to nonparametrically address the welfare-related issues listed in the introduction. In this respect, two important remarks are in order. A first remark relates to the efficient implementation (in computational terms) of the methodology. We believe that the IP and MILP formulations presented in the current paper convincingly show the computational tractability of the methodology. Still, we also believe that, in practice, considerable efficiency gains can be realized by exploiting the specificities of the collective rationality conditions; such enhancements of the computational efficiency can be particularly useful when there are many observations. For example, Cherchye, De Rock and Vermeulen (2005) suggest efficiency enhancing mechanisms that exploit a number of basic theoretical insights regarding the collective rationality tests; while these authors focus on the tests of collective rationality as they were originally proposed by Cherchye, De Rock and Vermeulen (2007), the same insights -and, probably, further refinements- are readily adapted to the (testing and recovery) methodology presented in this paper. Next, focusing on the specific IP/MILP structure of the proposed testing and recovery tools (with binary variables as the integer variables), one can conceive efficient solution algorithms that are specially tailored for addressing the testing and recovery questions that are relevant for the collective model; see, for example, the general discussion in Nemhauser and Wolsey (1999) on efficiently solving IP and MILP problems. Generally, we believe the development of efficient testing and recovery algorithms constitutes an interesting avenue for follow-up research.

The second remark that is relevant in view of practical applications pertains to the 'power' of the methodology that is proposed; this refers both to the probability of nonparametrically detecting violations of collective rationality by means of the testing tools, and to the possibility of providing tight bounds (on feasible income shares, personalized prices and personalized quantities) by means of the recovery tools. (See Andreoni and Harbaugh (2006) for a recent discussion of the power of revealed preference tests and a survey of nonparametric power assessment tools that are currently available.) While we have illustrated that the methodology can yield powerful (recovery) results even if there is no assignable quantity information, we have also argued that in general we can expect the power to increase (often substantially) when more assignable quantity information is available. In our opinion, this pleads for investing in collective consumption data sets that incorporate such information; such detailed data sets seem specially valuable for nonparametric welfare analysis in terms of the collective consumption model. Next, the power of the nonparametric methodology can be further increased by adapting the 'sequential maximum power path' idea of Blundell, Browning and Crawford (2003, 2006), who originally focused on the $GARP$ condition for the unitary model. Essentially, the approach of Blundell, Browning and Crawford uses estimated Engel curves for given price regimes to construct 'virtual' quantity bundles that maximize the power of the nonparametric (testing and recovery) tools. When adapting this approach to the methodology presented in this paper, one can focus on Engel curves for assignable quantities as well as on Engel curves for the aggregate household quantities. In our opinion, such extensions can be particularly valuable in view of real-life empirical applications.

Appendix: proofs

Proof of Proposition 1

Varian (1982) has proven equivalence between conditions (ii) and (iii), so we can restrict to proving equivalence between conditions (i) and (iii). This proof extends the proof of Proposition 1 of Cherchye, De Rock and Vermeulen (2007), who consider two-member households and do not account for the possibility of assignable quantity information.

1. Necessity. Under condition (i), we have that each $\widehat{\mathbf{q}}_t = \left(\mathfrak{q}_t^1, ..., \mathfrak{q}_t^M, \mathfrak{Q}_t\right)$ solves the problem

$$\max_{(\mathfrak{q}^1,...,\mathfrak{q}^M,\mathfrak{Q})} \sum_{m=1}^{M} \mu_t^m U^m \left(\mathfrak{q}^m, \mathfrak{Q}\right) \text{ s.t. } \mathbf{p}_t'[\sum\nolimits_{m=1}^{M} \mathfrak{q}^m + (\sum\nolimits_{m=1}^{M} \mathfrak{Q}^m + \mathfrak{Q}^h)] \leq \mathbf{p}_t'\mathbf{q}_t \text{ and } \mathfrak{q}^m \geq \mathbf{q}_t^{Am}.$$

Given concavity, the functions U^m are subdifferentiable, which carries over to their weighted sum $\sum_{m=1}^{M} \mu_t^m U^m$.[8] An optimal solution to the above maximization problem must therefore satisfy (for η_t the Lagrange multiplier associated with the budget constraint)

$$\mu_t^m U_{\mathfrak{q}_t^m}^m \leq \eta_t \mathbf{p}_t \text{ and } \sum_{m=1}^{M} \mu_t^m U_{\mathbf{Q}_t^c}^m \leq \eta_t \mathbf{p}_t,$$

for $U_{\mathfrak{q}_t^m}^m$ a subgradient of the function U^m defined for the vector $\mathfrak{q}^m$ and evaluated at $\mathfrak{q}_t^m$, and $U_{\mathbf{Q}_t^c}^m$ a subgradient defined for $\mathfrak{Q}^c$ and evaluated at $\mathfrak{Q}_t^c$ $(c = 1, ..., M, h)$. Letting $\mathfrak{P}_t^{m,c} = \frac{U_{\mathbf{Q}_t^c}^m}{\eta_t}$ and $\lambda_t^m = \frac{\eta_t}{\mu_t^m}$ thus gives for each m

$$U_{\mathfrak{q}_t^m}^m \leq \lambda_t^m \mathbf{p}_t \text{ and } U_{\mathbf{Q}_t^c}^m \leq \lambda_t^m \mathfrak{P}_t^{m,c}. \tag{6.1}$$

Next, concavity of the functions U^m implies for each m

$$U^m \left(\widehat{\mathbf{q}}_s\right) - U^m \left(\widehat{\mathbf{q}}_t\right) \leq U_{\mathfrak{q}_t^m}^m \left(\mathfrak{q}_s^m - \mathfrak{q}_t^m\right) + \sum_{c=1,...,M,h} U_{\mathbf{Q}_t^c}^m \left(\mathfrak{Q}_s^c - \mathfrak{Q}_t^c\right). \tag{6.2}$$

Substituting (6.1) in (6.2) and setting $U_k^m = U^m \left(\widehat{\mathbf{q}}_k\right)$ $(k = s, t)$ obtains condition (iii) of the proposition.

2. Sufficiency. Under condition (iii), for any $\widehat{\mathbf{q}} = \left(\mathfrak{q}^1, ..., \mathfrak{q}^M, \mathfrak{Q}\right)$ such that $\mathbf{p}_t'[\sum_{m=1}^{M} \mathfrak{q}^m +$

[8] To be precise, $-U^m$ $(m = 1, ..., M)$ is convex and therefore subdifferentiable. This, of course, does not affect our argument.

$(\sum_{m=1}^{M}\mathfrak{Q}^m+\mathfrak{Q}^h)] \leq \mathbf{p}_t'\mathbf{q}_t$ and $\mathfrak{q}^m \geq \mathbf{q}_t^{Am}$ we can define for all m

$$U^m(\widehat{\mathbf{q}}) = \min_{t\in\{1,...,T\}} \left[U_t^m + \lambda_t^m (\widehat{\mathbf{p}}_t^m)'(\widehat{\mathbf{q}}-\widehat{\mathbf{q}}_t)\right]. \tag{6.3}$$

Varian (1982) proves that $U^m(\widehat{\mathbf{q}}_t) = U_t^m$. Next, given $\mu_j^m \in \Re_{++}$, we have that

$$\sum_{m=1}^{M} \mu_t^m U^m(\widehat{\mathbf{q}}) \leq \sum_{m=1}^{M} \mu_t^m \left[U_t^m + \lambda_t^m (\widehat{\mathbf{p}}_t^m)'(\widehat{\mathbf{q}}-\widehat{\mathbf{q}}_t)\right].$$

Without losing generality, we concentrate on $\mu_t^m = (1/\lambda_t^m)$, which obtains

$$\sum_{m=1}^{M} \mu_t^m U^m(\widehat{\mathbf{q}}) \leq \sum_{m=1}^{M} \mu_t^m U_t^m + (\mathbf{p}_t)'(\mathbf{q}-\mathbf{q}_t),$$

for $\mathbf{q} = [\sum_{m=1}^{M}\mathfrak{q}^m + (\sum_{m=1}^{M}\mathfrak{Q}^m+\mathfrak{Q}^h)]$.

Since $\mathbf{p}_t'\mathbf{q} \leq \mathbf{p}_t'\mathbf{q}_t$, we thus have

$$\sum_{m=1}^{M} \mu_t^m U^m(\widehat{\mathbf{q}}) \leq \sum_{m=1}^{M} \mu_t^m U_t^m = \sum_{m=1}^{M} \mu_t^m U_t^m(\widehat{\mathbf{q}}_t),$$

which proves that $\widehat{\mathbf{q}}_t$ maximizes $\sum_{m=1}^{M} \mu_t^m U^m(\mathfrak{q}^m, \mathfrak{Q})$ subject to $\mathbf{p}_t'[\sum_{m=1}^{M}\mathfrak{q}^m+(\sum_{m=1}^{M}\mathfrak{Q}^m+\mathfrak{Q}^h)] \leq \mathbf{p}_t'\mathbf{q}_t$ and $\mathfrak{q}^m \geq \mathbf{q}_t^{Am}$. We conclude that the functions U^m in (6.3) provide a collective rationalization of the set S^A. These functions have the properties listed in condition (i) of the proposition (compare with Varian, 1982).

Proof of Proposition 2

1. Necessity. Suppose there exist feasible personalized prices $(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M)$ such that for each member $m = 1, ..., M$ the set $\{(\mathfrak{P}_t^{m,h}; \mathbf{q}_t); t = 1, ..., T\}$ satisfies *GARP*. Then the corresponding specification of $\mathfrak{P}_t^{m,h}$, $\widehat{y}_t^m$ and $x_{st}^m \in \{0,1\}$ satisfies rules (PP-i)-(PP-v). First, rules (PP-i) and (PP-ii) are satisfied because the feasible personalized prices and income shares are consistent with Definitions 3 and 5. Next, to see consistency with rules (PP-iii)-(PP-v), consider any sequence $(u, v, ..., z)$ such that $(\mathfrak{P}_s^{m,h})'\mathbf{q}_s \geq (\mathfrak{P}_s^{m,h})'\mathbf{q}_u$, $(\mathfrak{P}_u^{m,h})'\mathbf{q}_u \geq (\mathfrak{P}_u^{m,h})'\mathbf{q}_v$, ..., $(\mathfrak{P}_z^{m,h})'\mathbf{q}_z \geq (\mathfrak{P}_z^{m,h})'\mathbf{q}_t$. (Trivially, another sequence does not impose restrictions on x_{st}^m.) Rule (PP-iii) then implies $x_{su}^m = x_{uv}^m = ... = x_{zt}^m = 1$, and rule (PP-iv) consequently obtains $x_{st}^m = 1$. Rule (PP-v) is then automatically satisfied because the set $\{(\mathfrak{P}_t^{m,h}; \mathbf{q}_t); t = 1, ..., T\}$ satisfies *GARP*, and thus $(\mathfrak{P}_t^{m,h})'\mathbf{q}_t \leq (\mathfrak{P}_t^{m,h})'\mathbf{q}_s$ whenever $(\mathfrak{P}_s^{m,h})'\mathbf{q}_s \geq (\mathfrak{P}_s^{m,h})'\mathbf{q}_u$, $(\mathfrak{P}_u^{m,h})'\mathbf{q}_u \geq (\mathfrak{P}_u^{m,h})'\mathbf{q}_v$, ..., $(\mathfrak{P}_z^{m,h})'\mathbf{q}_z \geq (\mathfrak{P}_z^{m,h})'\mathbf{q}_t$ (which corresponds to $\widehat{\mathbf{q}}_s\ R^m\ \widehat{\mathbf{q}}_t$).

2. Sufficiency. If there exist $\mathfrak{P}_t^{m,h}$, $\widehat{y}_t^m$ and $x_{st}^m \in \{0,1\}$ that satisfy rules (PP-i)-(PP-v), then there exist feasible personalized prices $(\widehat{\mathbf{p}}_t^1, ..., \widehat{\mathbf{p}}_t^M)$ such that for each member $m = 1, ..., M$

the set $\{\left(\mathfrak{P}_t^{m,h};\mathbf{q}_t\right);\, t=1,...,T\}$ satisfies *GARP*. We prove *ex absurdo*. Suppose that for any specification of $\left(\widehat{\mathbf{p}}_t^1,...,\widehat{\mathbf{p}}_t^M\right)$ we have a sequence $(u,\, v,\, ...,\, z)$ such that $\widehat{\mathbf{q}}_s\ R_0^m\ \widehat{\mathbf{q}}_u$, $\widehat{\mathbf{q}}_u\ R_0^m\ \widehat{\mathbf{q}}_v$, ..., $\widehat{\mathbf{q}}_z\ R_0^m\ \widehat{\mathbf{q}}_t$ and $\left(\mathfrak{P}_t^{m,h}\right)'\mathbf{q}_t > \left(\mathfrak{P}_t^{m,h}\right)'\mathbf{q}_s$. By construction, $\widehat{\mathbf{q}}_s\ R_0^m\ \widehat{\mathbf{q}}_u$, $\widehat{\mathbf{q}}_u\ R_0^m\ \widehat{\mathbf{q}}_v$, ..., $\widehat{\mathbf{q}}_z\ R_0^m\ \widehat{\mathbf{q}}_t$ implies $\left(\mathfrak{P}_s^{m,h}\right)'\mathbf{q}_s \geq \left(\mathfrak{P}_s^{m,h}\right)'\mathbf{q}_u$, $\left(\mathfrak{P}_u^{m,h}\right)'\mathbf{q}_u \geq \left(\mathfrak{P}_u^{m,h}\right)'\mathbf{q}_v$, ..., $\left(\mathfrak{P}_z^{m,h}\right)'\mathbf{q}_z \geq \left(\mathfrak{P}_z^{m,h}\right)'\mathbf{q}_t$. In terms of the rules (PP-i)-(PP-v), this means that there always exists a sequence $(u,\, v,\, ...,\, z)$ such that, on the one hand, $x_{su}^m = x_{uv}^m = ... = x_{zt}^m = 1$ (because of (PP-iii)) and thus $x_{zt}^m = 1$ (because of (PP-iv)) while, on the other hand, $\left(\mathfrak{P}_t^{m,h}\right)'\mathbf{q}_t > \left(\mathfrak{P}_t^{m,h}\right)'\mathbf{q}_s$ and thus rule (PP-v) is violated. In other words, there does not exist $\mathfrak{P}_t^{m,h}$, $\widehat{y}_t^m$ and $x_{st}^m \in \{0,1\}$ that simultaneously satisfy rules (PP-i)-(PP-v).

Proof of Proposition 3

The proof is directly analogous to that of Proposition 2.

Proof of Lemma 1

Rule (i):

1. Necessity. If for all sets $\widehat{S}^A$ there exists m such that $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$, then $\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\mathbf{q}_t$. We prove *ex absurdo*. Suppose for all sets $\widehat{S}^A$ there exists m such that $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$, and $\mathbf{p}_s'\mathbf{q}_s < \mathbf{p}_s'\mathbf{q}_t$. This is impossible because $\mathbf{p}_s'\mathbf{q}_s < \mathbf{p}_s'\mathbf{q}_t$ implies there exists at least one set $\widehat{S}^A$ such that not $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$ for all m. More specifically with an $\widehat{S}^A$ such that $\left(\widehat{\mathbf{p}}_s^m\right)\widehat{\mathbf{q}}_s = \left(\mathbf{p}_s'\mathbf{q}_s\right)/M$ and $\left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_t = \left(\mathbf{p}_s'\mathbf{q}_t\right)/M$.

2. Sufficiency. Recall that each $\widehat{\mathbf{q}}_t = \left(\mathfrak{q}_t^1,...,\mathfrak{q}_t^M,\mathfrak{Q}_t\right)$ satisfies $\mathbf{q}_t = \sum_{m=1}^M \mathfrak{q}_t^m + \left(\sum_{m=1}^M \mathfrak{Q}_t^m + \mathfrak{Q}_t^h\right)$, and each $\widehat{\mathbf{p}}_t^m = \left(\mathfrak{p}_t^{m,1},...,\mathfrak{p}_t^{m,M},\mathfrak{P}_t^m\right)$ satisfies $\mathfrak{p}_t^{m,m} = \mathbf{p}_t$, $\mathfrak{p}_t^{m,l} = \mathbf{0}$ for $l \neq m$ and $\mathfrak{P}_t^m = \left(\mathfrak{P}_t^{m,1},...,\mathfrak{P}_t^{m,M},\mathfrak{P}_t^{m,h}\right)$ such that $\mathbf{p}_t = \sum_{m=1}^M \mathfrak{P}_t^{m,c}$ for all c. Given this, $\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\mathbf{q}_t$ implies for any $\widehat{S}^A$ there exists m such that $\left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_s \geq \left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_t$ and thus $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$.

Rule (ii):

1. Necessity. $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$ for all sets $\widehat{S}^A$ implies $\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right)$. We prove *ex absurdo*. Suppose $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$ for all sets $\widehat{S}^A$ and $\mathbf{p}_s'\mathbf{q}_s^{Am} < \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right)$. This is impossible because $\mathbf{p}_s'\mathbf{q}_s^{Am} < \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right)$ implies there exists at least one set $\widehat{S}^A$ with not $\widehat{\mathbf{q}}_s R_0^m \widehat{\mathbf{q}}_t$. More specifically with an $\widehat{S}^A$ such that $\left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_s = \mathbf{p}_s'\mathbf{q}_s^{Am}$ and $\left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_t = \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right)$.

2. Sufficiency. By construction, $\mathbf{p}_s'\mathbf{q}_s^{Am} \geq \mathbf{p}_s'\left(\mathbf{q}_t - \sum_{l=1,l\neq m}^M \mathbf{q}_t^{Al}\right)$ implies $\left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_s \geq \left(\widehat{\mathbf{p}}_s^m\right)'\widehat{\mathbf{q}}_t$ for any $\widehat{S}^A$, and thus $\widehat{\mathbf{q}}_s R_o^m \widehat{\mathbf{q}}_t$ for any $\widehat{S}^A$.

Proof of Proposition 4

Given that a collective rationalization of the set S^A is possible, we consider a set $\widehat{S}^A$ that is consistent with condition (ii) in Proposition 1. Using Definition 4, this set $\widehat{S}^A$ defines relations R_0^m and R^m $(m = 1, 2)$. We will show that these relations (defined in terms of feasible personalized quantities $\widehat{\mathbf{q}}_t$) must satisfy the analogues of rules (i)-(vii) in Proposition 4. These requirements carry over to the hypothetical relations H_0^m and H^m (defined in terms of observed quantities $\mathbf{q}_t$) specified in Proposition 4: a collective rationalization of the set S^A is possible only if there exists at least one specification of these hypothetical relations that is consistent with these requirements.

Rules (i) and (ii): These rules follow directly from Lemma 1.

Rule (iii): This rule imposes transitivity.

Rule (iv): For all $m \in \mathbf{M}$ we have $\widehat{\mathbf{q}}_{t_{k(m)}}$ R^m $\widehat{\mathbf{q}}_s$ for $k(m) \leq M^*$, which requires $(\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s \leq (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_{k(m)}}$ because of condition (ii) in Proposition 1. As a result we have $\sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s \leq \sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_{k(m)}}$. Next, recall the definitional fact $\sum_{m=1}^{M} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s = \mathbf{p}_s' \mathbf{q}_s$. Given this, $\mathbf{p}_s' \mathbf{q}_s \geq \mathbf{p}_s' \left(\sum_{k=1}^{M^*} \mathbf{q}_{t_k} \right)$ necessarily implies $\sum_{l \notin \mathbf{M}} (\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_s \geq \left(\mathbf{p}_s' \left(\sum_{k=1}^{M^*} \mathbf{q}_{t_k} \right) - \sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_{k(m)}} \right)$. Thus, because $\mathbf{p}_s' \mathbf{q}_{t_k} = \sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_k} + \sum_{l \notin \mathbf{M}} (\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_{t_k}$ for all $k \leq M^*$, there must exist $l \notin \mathbf{M}$ and $k \leq M^*$ such that $(\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_s \geq (\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_{t_k}$, or $\widehat{\mathbf{q}}_s$ R_0^l $\widehat{\mathbf{q}}_{t_k}$.

Rule (v): For all $m \in \mathbf{M}$ we have $\widehat{\mathbf{q}}_{t_{k(m)}}$ R^m $\widehat{\mathbf{q}}_s$ for $k(m) \leq M^* - 1$, which requires $(\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s \leq (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_{k(m)}}$. As a result we have $\sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s \leq \sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_{k(m)}}$. Next, by construction $\sum_{m=1}^{M} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_s = \mathbf{p}_s' \mathbf{q}_s$ and $\sum_{m \in \mathbf{M}} (\widehat{\mathbf{p}}_s^m)' \widehat{\mathbf{q}}_{t_{k(m)}} \leq \mathbf{p}_s' \left(\sum_{k=1}^{M^*-1} \mathbf{q}_{t_k} \right)$. Given this, $\mathbf{p}_s' \mathbf{q}_s \geq \mathbf{p}_s' \left(\sum_{k=1}^{M^*} \mathbf{q}_{t_k} \right)$ necessarily implies $\sum_{l \notin \mathbf{M}} (\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_s \geq \mathbf{p}_s' \mathbf{q}_{t_{M^*}}$. In turn, this implies that there must exist $l \notin \mathbf{M}$ such that $(\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_s \geq (\widehat{\mathbf{p}}_s^l)' \widehat{\mathbf{q}}_{t_{M^*}}$, or $\widehat{\mathbf{q}}_s$ R_0^l $\widehat{\mathbf{q}}_{t_{M^*}}$.

Rule (vi): For all m we have $\widehat{\mathbf{q}}_{s_{k(m)}}$ R^m $\widehat{\mathbf{q}}_t$ for some $k(m) \leq M^*$, which requires $(\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t \leq (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_{s_{k(m)}}$. As result we have $\sum_{m=1}^{M} (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t \leq \sum_{m=1}^{M} (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_{s_{k(m)}}$. Next, by construction we have $\sum_{m=1}^{M} (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t = \mathbf{p}_s' \mathbf{q}_s$ and $\sum_{m=1}^{M} (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_{s_{k(m)}} \leq \mathbf{p}_t' \left(\sum_{k=1}^{M^*} \mathbf{q}_{s_k} \right)$. As a result, we obtain the requirement $\mathbf{p}_t' \mathbf{q}_t \leq \sum_{k=1}^{M^*} \mathbf{p}_t' \mathbf{q}_{s_k}$.

Rule (vii): We have $\widehat{\mathbf{q}}_s$ R^m $\widehat{\mathbf{q}}_t$, which requires $(\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t \leq (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_s$. By construction, we have $(\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_t \geq \mathbf{p}_t' \mathbf{q}_t^{Am}$ and $\mathbf{p}_t' \left(\mathbf{q}_s - \sum_{l=1, l \neq m}^{M} \mathbf{q}_s^{Al} \right) \geq (\widehat{\mathbf{p}}_t^m)' \widehat{\mathbf{q}}_s$. Hence, we must have $\mathbf{p}_t' \mathbf{q}_t^{Am} \leq \mathbf{p}_t' \left(\mathbf{q}_s - \sum_{l=1, l \neq m}^{M} \mathbf{q}_s^{Al} \right)$.

Proof of Proposition 5

The proof is directly analogous to that of Proposition 2 (necessity part).

References

[1] Afriat, S. (1967), "The construction of utility functions from expenditure data", *International Economic Review* 8, 67-77.

[2] Andreoni, J. and W.T. Harbaugh (2006), "Power indices for revealed preference tests", *University of Wisconsin-Madison Department of Economics Working Paper* 2005-10, version 2006.

[3] Blow, L. and I.A. Crawford (2001), "The cost of living with the RPI: Substitution bias in the UK Retail Prices Index", *Economic Journal* 111, 311-334.

[4] Blundell, R.W, M. Browning and I.A. Crawford (2003), "Nonparametric Engel curves and revealed preferences", *Econometrica* 71, 205-240.

[5] Blundell, R.W, M. Browning and I.A. Crawford (2006), "Best nonparametric bounds on demand responses", *CeMMAP Working Papers* CWP 12/05, version 2006.

[6] Blundell, R., P.-A. Chiappori and C. Meghir (2005), "Collective labor supply with children", *Journal of Political Economy* 113, 1277-1306.

[7] Bourguignon, F., Browning, M and P.-A. Chiappori (2006), "Efficient intra-household allocations and distribution factors: implications and identification", *CAM Working Paper* 2006-02.

[8] Browning, M, F. Bourguignon, P.-A. Chiappori and V. Lechene (1994), "Income and Outcomes: A Structural Model of Intrahousehold Allocations", *Journal of Political Economy* 102, 1067-1096.

[9] Browning, M. and P.-A. Chiappori (1998), "Efficient intra-household allocations: a general characterization and empirical tests", *Econometrica* 66, 1241-1278.

[10] Browning, M., P.-A. Chiappori and A. Lewbel (2006), "Estimating consumption economies of scale, adult equivalence scales, and household bargaining power", *Boston College Working Paper* WP588, Department of Economics, Boston College, version 2006.

[11] Cherchye, L., B. De Rock and F. Vermeulen (2007), "The collective model of household consumption: a nonparametric characterization", *Econometrica* 75, 553-574.

[12] Cherchye, L., B. De Rock and F. Vermeulen (2005), "Opening the black box of intra-household decision-making: theory and non-parametric empirical tests of general collective consumption models", *IZA Discussion Paper* No. 1603; *CentER Discussion Paper* No. 2005-51.

[13] Chiappori, P.-A. (1988), "Rational household labor supply", *Econometrica* 56, 63-89.

[14] Chiappori, P.-A. (1992), "Collective labor supply and welfare", *Journal of Political Economy* 100, 437-467.

[15] Chiappori, P.-A. and I. Ekeland (2005): "The micro economics of group behavior: identification", *University of Chicago Working Paper.*

[16] Chiappori, P.-A. and I. Ekeland (2006), "The micro economics of group behavior: general characterization", *Journal of Economic Theory* 130, 1-26.

[17] Michaud, P.-C. and F. Vermeulen (2006), "A collective labor supply model: identification and estimation in the presence of externalities by means of panel data", *RAND Labor and Population Program* WR-406, Santa Monica, RAND Corporation.

[18] Nemhauser, G.L. and L.A. Wolsey (1999), *Integer and combinatorial optimization*, John Wiley, New York.

[19] Sethuraman, J., T.C. Piaw and R.V. Vohra (2003), "Integer programming and Arrovian Social Welfare Functions", *Mathematics of Operations Research* 28, 309-326.

[20] Varian, H.R. (1982), "The nonparametric approach to demand analysis", *Econometrica* 50, 945-972.

[21] Varian, H.R. (2006), "Revealed preference", in M. Szenberg, L.Ramrattan and A.A. Gottesman (eds.), *Samuelsonian economics and the 21st century*, Oxford University Press.